1999

Tracing the development of the inverse base-rate effect in category learning.

Ferne Joi Friedman-berg
University of Massachusetts Amherst

TRACING THE DEVELOPMENT OF THE INVERSE BASE-RATE EFFECT IN CATEGORY LEARNING

A Thesis Presented

by

FERNE JOI FRIEDMAN-BERG

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE

February 1999

Psychology

TRACING THE DEVELOPMENT OF THE INVERSE BASE-RATE EFFECT IN CATEGORY LEARNING

A Thesis Presented

by

FERNE JOI FRIEDMAN-BERG

Approved as to style and content by:

Arnold D. Well, Chair

James I. Chumbley, Member

Marvin W. Daehler, Member

Melinda A. Novak, Department Head
Department of Psychology

ACKNOWLEDGEMENTS

I would like to thank a number of people who have seen me through the writing of this thesis. First, I would like to thank Arnold Well for his time and patience. He took the time to show me how to organize my thoughts into a clear and cohesive argument and had the patience to read through numerous drafts of this thesis on the way to its completion. I would also like to thank my parents, Arlene and Henry Friedman, who have always supported me and had faith in me even when times were rough. Finally, I would most like to thank my husband Ed and my son Grant. I couldn't have done it without their love and support.

ABSTRACT

TRACING THE DEVELOPMENT OF THE INVERSE BASE-RATE EFFECT IN CATEGORY LEARNING

FEBRUARY 1999

FERNE JOI FRIEDMAN-BERG, B.A., TEMPLE UNIVERSITY

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Arnold D. Well

In this thesis we examined how the inverse base-rate effect (Medin & Edelson, 1988) develops over time and explored the associations of category features to categories formed during learning which may cause the effect. Six experiments are presented which extend the findings of previous research into the inverse base-rate effect. In these experiments, after first establishing an inverse base-rate effect, we manipulated frequencies using a simplified version of Medin & Bettger's (1991) Experiment 2 and examined the separate responses of fast and slow learners. When testing for the inverse base-rate effect throughout learning while manipulating frequencies we found that the development of the inverse base-rate effect is intrinsically tied to the frequencies which are present during learning. The effect appeared to develop whether those unequal category frequencies are present early or later in learning, contingent upon when learning takes place. Also found was an inverse triplet base-rate effect. Finally, we examined a proposal made by Kruschke (1996) in which he states that the inverse base-rate effect is caused by the perfectly predictive feature(s) of a rare category being learned in contrast to

the perfectly predictive feature(s) of a common category. When conditional probability estimates were examined we could find no direct evidence for his proposal.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS..... iii

ABSTRACT..... iv

LIST OF TABLES..... ix

Chapter

1. FREQUENCY MANIPULATIONS AND CATEGORY LEARNING..... 1

Introduction..... 1
Base-rate Neglect..... 2
The Inverse Base-rate Effect..... 4
Questions that Remain About the Inverse Base-rate Effect..... 9

2. MODELING THE EFFECTS OF FREQUENCY CHANGES ON CATEGORY LEARNING..... 12

Exemplar Models..... 12

The Context Model..... 12
The Modified Context Model..... 14
Concept Learning by Exemplar Memorization..... 15

Connectionist Models..... 17

The Adaptive Network Model..... 17
The Modified Adaptive Network Model..... 19
The Distributed Stimulus Sampling Model..... 21
The Attentional Connectionist Model..... 23
The Attention to Distributed InpuT Model (ADIT)..... 24

3. PRELIMINARY RESEARCH-EXPERIMENTS 1 AND 2..... 29

Experiment 1..... 29

Method..... 30

Participants..... 30
Stimuli and Procedure..... 30

Results and Discussion..... 32

Experiment 2..... 35

Method..... 37

Participants..... 37
Stimuli and Procedure..... 37

Results and Discussion..... 38

4. EXAMINING THE DEVELOPMENT OF THE INVERSE BASE-RATE EFFECT IN EARLY AND LATE LEARNERS..... 42

Experiments 3 and 4..... 42

Method..... 44

Participants..... 44
Stimuli and Procedure..... 44

Learning Data..... 46
Test Data..... 48

The Inverse Base-rate Effect..... 48
The Inverse Base-rate Effect in Early and Late Learners... 49
The I+PC+PR Triplet..... 50
The I+PC+PR Triplet in Early and Late Learners..... 51
The Imperfect Predictor..... 53
The Perfect Common and Perfect Rare Predictors..... 55
I+PC and I+PR..... 56

Discussion..... 57

5. CONDITIONAL PROBABILITY ESTIMATES AS A MEASURE OF WHAT IS KNOWN ABOUT CATEGORY FEATURES..... 60

Experiments 5 and 6..... 60

Method..... 62

Participants..... 62
Stimuli and Procedure..... 62

Results-Learning and Conditional Probability Estimates 65

Learning Data..... 65
Asymmetries in Estimates for P(S|D)
And P(D|S)..... 67
Summary-Overall Estimation Data..... 72
Asymmetries in Estimates for P(S|D)
and P(D|S) for Early and Late Learners..... 73
Summary-Estimation Data for Early and Late Learners..... 77
Accuracy in Estimates for P(S|D) and P(D|S) for
Early and Late Learners..... 78

Results-Test Items..... 79

The Inverse Base-rate Effect..... 79
The Inverse Base-rate Effect in Early and Late Learners ... 80
The I+PC+PR Triplet..... 82
The I+PC+PR Triplet in Early and Late Learners..... 82
The Imperfect Predictor..... 85
The Perfect Common and Perfect Rare Predictors..... 87

Summary-Test Item Data..... 89

6. GENERAL DISCUSSION..... 91

Conclusion..... 91
Some Additional Questions for Future Research..... 94

APPENDICES

A. TABLES FOR EXPERIMENT 2..... 96
B. TABLES FOR EXPERIMENTS 3 AND 4..... 97
C. TABLES FOR EXPERIMENTS 5 AND 6..... 101

BIBLIOGRAPHY..... 114

LIST OF TABLES

Table		Page
1.	Category structure for Gluck & Bower Experiment 1 (1988b).....	4
2.	Disease and symptom structure for Medin & Edelson (1988) Experiment 1.....	5
3.	Test items and response proportions for Medin & Edelson (1988) Experiment 1.....	6
4.	Disease structure for Medin & Bettger (1991) Experiment 2.....	8
5.	Response proportions to test items from Medin & Bettger (1991) Experiment 2	8
6.	Disease and symptom structure for Experiment 1.....	30
7.	Learning trial for Experiment 1.....	31
8.	Test items for Experiment 1.....	32
9.	Proportion correct per block in Experiment 1.....	33
10.	Response proportions obtained in Experiment 1.....	34
11.	Response proportions-Kruschke (1996).....	34
12.	Disease structure for Experiment 2.....	37
13.	Proportion correct per block in Experiment 2.....	39
14.	Response proportions to test items from Medin & Bettger (1991) Experiment 2.....	40
15.	Response proportions to test items for Experiment 2.....	40
16.	Response to PC+PR for the unequal late condition.....	41
17.	Test items for Experiments 3 and 4.....	45
18.	Proportion correct per block in Experiment 3.....	47
19.	Proportion correct per block in Experiment 4.....	47

20. Response proportions to test item PC+PR..... 48

21. Response proportions to test item PC+PR for early and late learners..... 49

22. Response proportions to test item I+PC+PR..... 51

23. Response proportions to test item I+PC+PR for early and late learners... 52

24. Response proportions for test item I..... 54

25. Response proportions to I for early and late learners..... 54

26. Response proportions for test item PC..... 55

27. Response proportions to test item PC for early and late learners..... 55

28. Response proportions to test item PR..... 56

29. Response proportions to test item PR for early and late learners..... 56

30. Test items for Experiments 5 and 6..... 65

31. Proportion correct per block in Experiment 5..... 66

32. Proportion correct per block in Experiment 6..... 67

33. Estimates for P(S|D) for the imperfect predictor-Experiment 5..... 68

34. Estimates for P(S| D) for the imperfect predictor-Experiment 6..... 68

35. Estimates for P(S| D) for the perfect predictor-Experiment 5..... 69

36. Estimates for P(D|S) for the perfect predictor-Experiment 5..... 70

37. Estimates for P(S| D) for the perfect predictor-Experiment 6..... 71

38. Estimates for P(D|S) for the perfect predictor-Experiment 6..... 72

39. P(S(imperfect)|D) for the unequal early condition
early versus late learners-Experiments 5 and 6..... 74

40. P(S(perfect)|D) for the unequal early condition – early versus late learners-Experiments 5 and 6..... 75

41. P(D|S(perfect)) for the unequal early condition – early versus late learners- Experiments 5 and 6..... 77

42. P(D|S(imperfect)) for the unequal early condition – early versus late learners-Experiments 5 and 6..... 77

43. Response proportions to test item PC+PR-Experiment 5..... 80

44. Response proportions to test item PC+PR-Experiment 6..... 81

45. Response proportions to test item I+PC +PR for Experiment 5..... 84

46. Response proportions to test item I+PC+PR for Experiment 6..... 85

47. Response proportions to the imperfect predictor-Experiment 5..... 86

48. Response proportions to the imperfect predictor-Experiment 6..... 86

49. Response proportions to the perfect common predictor-Experiment 5..... 87

50. Response proportions to the perfect common predictor-Experiment 6..... 88

51. Response proportions to the perfect rare predictor-Experiment 5..... 88

52. Response proportions to the perfect rare predictor-Experiment 6..... 89

53. Complete response proportions for Experiment 2..... 96

54. Complete data for test item I-Experiments 3 and 4..... 97

55. Complete data for test item PC-Experiments 3 and 4..... 97

56. Complete data for test item PR-Experiments 3 and 4..... 98

57. Complete data for test item PC+PR-Experiments 3 and 4..... 98

58. Complete data for test item I+PC+PR-Experiments 3 and 4..... 99

59. Complete data for test item I+PC-Experiments 3 and 4..... 99

60. Complete data for test item I+PR-Experiments 3 and 4..... 100

61. Complete probability estimates for P(S| D)-Experiment 5..... 101

62. Complete probability estimates for P(S| D)-Experiment 6..... 101

63. Complete probability estimates for P(D|S)-Experiment 5..... 102

64. Complete probability estimates for P(D|S)-Experiment 6..... 102

65. Complete probability estimates for P(S|D)-
early versus late-Experiment 5..... 103

66. Complete probability estimates for P(S|D)-
early versus late-Experiment 6..... 104

67. Complete probability estimates for P(D|S)-
early versus late-Experiment 5..... 105

68. Complete probability estimates for P(D|S)-
early versus late-Experiment 6..... 106

69. Complete data for test item I-Experiment 5..... 107

70. Complete data for test item I-Experiment 6..... 107

71. Complete data for test item PC-Experiment 5..... 108

72. Complete data for test item PC-Experiment 6..... 108

73. Complete data for test item PR-Experiment 5..... 109

74. Complete data for test item PR-Experiment 6..... 109

75. Complete data for test item PC+PR-Experiment 5..... 110

76. Complete data for test item PC+PR-Experiment 6..... 110

77. Complete data for test item I+PC+PR-Experiment 111

78. Complete data for test item I+PC+PR-Experiment 6..... 111

79. Complete data for test item I+PC*-Experiment 5..... 112

80. Complete data for test item I+PR*-Experiment 5..... 112

81. Complete data for test item PC+PR*-Experiment 5..... 113

82. Complete data for test item I+PC+PR*-Experiment 5..... 113

CHAPTER 1

FREQUENCY MANIPULATIONS AND CATEGORY LEARNING

Introduction

In recent decades, researchers have developed several classes of models to account for category learning. These include connectionist models, which have their foundation in the associationist literature, exemplar models, grounded in theories of information processing, and hybrid models that combine aspects of both exemplar and connectionist models. These models have been able to account for a wide variety of findings from category learning and transfer tasks (see Estes, 1994 and Shanks, 1995 for a complete review). However, although these different classes of models are often viewed as competing theories[1], it has been difficult to discriminate among exemplar and connectionist models. Even though these models appear to have quite different architectures and learning mechanisms, they are able to account for much of the same data.

One manipulation that seems to hold some promise for discriminating among different classes of models is to vary the frequencies with which features and categories are presented (Gluck & Bower,1988; Medin & Edelson,1988; Estes, Campbell, Hastopoulis, & Hurwitz,1989; Medin & Bettger, 1991; Myers, Lohmeir, & Well,1994).

In this thesis, after first briefly reviewing work that suggests that people often underutilize frequency information, we focus on the "inverse base-rate effect", a paradoxical finding that seems to suggest that under certain conditions, less frequently presented features can elicit stronger responses than more frequently presented ones. After considering some attempts to establish and model the effect, we report several studies aimed at increasing our understanding of how the inverse base-rate effect develops during category learning. We do this by presenting test trials at various stages of learning and by considering the performance of both fast and slow learners.

Base-rate Neglect

One line of research that holds promise for discriminating between connectionist and exemplar models examines how people learn and apply their knowledge of base-rate information when the frequencies with which exemplars and categories occur in a category learning task are varied. There is a long history of research that shows people are poor in utilizing base-rate information when performing judgment tasks (e.g., Kahneman & Tversky, 1973; but see Koehler, 1996). Although base-rate information is more likely to be used when exemplars of categories are presented trial-by-trial at the appropriate relative frequencies (Christensen-Szelanski & Bushyhead, 1981; Christensen-Szelanski & Beach, 1982; Koehler, 1996), than when base rates are given in numerical

[1]Although most researchers see exemplar models and connectionist models as competing theories, there are a few who say that they are compatible, and only provide different levels of description (Broadbent, 1985; Shanks, 1995). We will take the more common stance that they are competing types of models.

statements and not experienced directly, base-rate information is still is underused, and often misused.

One example of the underuse of base-rate information in category learning is the base-rate neglect described by Gluck & Bower (1988a and 1988b, see also Estes, Campbell, Hastopoulis, & Hurwitz, 1989; and Nosofsky, Kruschke, & McKinley, 1992). In Gluck & Bower's (1988b) Experiment 1, for example, participants had to learn about two diseases, C (common) and R (rare), that occurred in a 3:1 ratio. For the rare disease, symptoms 1, 2, 3, and 4 occurred with probabilities .6, .4, .3, and .2 respectively. For the common disease, the symptom probabilities were reversed, so the objective probability for either the common or rare disease, given symptom 1, was equal to .5 (see Table 1). Participants first received a series of learning trials on which they were given feedback, then received test trials without feedback. On some of the test trials, participants were presented with a single symptom and were asked to indicate the probability of the rare disease. Although the objective probability of the rare disease, given symptom 1 was .5 if one takes account of the base-rate information, the average estimate given was .67.[2]

Gluck & Bower (1988a and 1988b) were able to show that a connectionist model, using the Least Mean Squares rule to change the connection weights in an associative network, could account for the results. Nosofsky et al. (1992) and Gluck & Bower (1988b) demonstrated that the typical exemplar model predicted normative use of base-rate information, not the base-rate neglect shown by participants. However, more recent work (Myers, Lohmeier, & Well, 1994) has shown that exemplar models that incorporate

error-driven learning (e.g., Chumbley, 1986), are also able to predict base-rate neglect. Although this is good news for researchers who favor exemplar models, it still leaves us in a position where we are unable to discriminate between the two types of models.

Table 1. Category structure for Gluck & Bower Experiment 1 (1988b)

Symptom (S)	RARE (R)	COMMON (C)	P(R | S)[3]
1	.6	.2	.50
2	.4	.3	.31
3	.3	.4	.20
4	.2	.6	.10

The Inverse Base-rate Effect

Medin & Edelson (1988) investigated a phenomenon referred to as the inverse base-rate effect (described earlier as the "novelty" effect by Binder & Estes, 1966). In a series of classification experiments, participants were found in some instances to use base-rate information correctly, but in others to seemingly ignore or underweight the base rates.

In Medin & Edelson's (1988) Experiment 1, participants first learned disease and symptom pairings with feedback given, then received a series of test trials with no feedback. Within each problem structure, there were three symptoms and two diseases.

---

[2] In Gluck and Bower (1988) participants gave numerical estimates but other researchers have used a forced choice paradigm (see Estes et al., 1989, and Nosofsky et al., 1992). Results from the forced choice paradigm were not distinguishable from the probability estimation results.

[3] Bayesian probability $P(R \mid S)$ is given by: $$P(R \mid S) = \frac{P(S|R)P(R)}{P(S|R)P(R)+P(S|C)P(C)}$$

Each disease had two symptoms, and one disease (i.e., the common disease, C) occurred three times as often as the rare disease (R). Both diseases in the pair shared a symptom (the imperfect predictor, I), and each disease also had one symptom that was a perfect predictor (PC or PR, see Table 2 for general disease structure).

Participants went through a series of up to 192 learning trials with three problem structures (using 9 symptoms and 6 diseases) and then were tested with novel combinations of symptoms in a transfer task. On some test trials, participants were asked to predict the disease that went with I, PC + PR, or I + PC + PR (see Table 3). When participants were given the imperfect predictor (I) by itself, they were more likely to predict the common disease, consistent with the conditional probabilities of the common and rare diseases given the imperfect predictor. They also responded according to the disease base rates when given a triplet consisting of the imperfect predictor, the perfect predictor of the common disease, and the perfect predictor of the rare disease (Medin & Edelson, 1988).

Table 2. Disease and symptom structure
for Medin & Edelson (1988) Experiment 1

Disease	Symptoms	Frequency
C	I + PC	3
R	I + PR	1

Note: participants received 3 pairs of diseases with this structure

where R is the rare disease, C is the common disease and S is the symptom.

Table 3. Test items and response proportions for Medin & Edelson (1988) Experiment 1

Symptoms	Proportion of Responses: Common Disease	Proportion of Responses: Rare Disease
I	.781	.146
PC	.812	—
PR	—	.927
PC + PR	.323	.584
I+PC+PR	.708	.281

Note: proportions do not add to 1.00 because error responses (responses to a disease in a non-relevant disease pair) are not included in the table.

The critical finding of Medin & Edelson (1988) was that when participants were presented with the combination of the perfect common symptom and the perfect rare symptom (PC + PR) they were more likely to respond with the rare disease. It appears reasonable to expect that when presented with PC + PR, participants should have chosen the common disease because, according to the odds form of Bayes' Theorem:

$$\frac{P(C|PR+PC)}{P(R|PR+PC)} = \frac{P(PR+PC|C)P(C)}{P(PR+PC|R)P(R)} \qquad [1]$$

If participants initially had no reason to expect that P(PC+PR|C) would differ from P(PC+PR|R) (Shanks, 1992), then, because the P(C)= 3P(PR), the ratio of P(C|PC+PR) to P(R|PC+PR) becomes 3 to 1 which would lead participants to choose the common disease. However, they chose the rare disease more frequently (see Table 3). This paradoxical finding was called the inverse base-rate effect.

Medin & Bettger (1991) proposed that the inverse base-rate effect is driven primarily by early learning. They provided evidence for the primacy of early learning in their second experiment in which four different category pairs were studied. Each category pair was presented at different frequency ratios over the four blocks (see Table 4). Within the four pairs, the two pairs that would give the best contrast for understanding the differential effect of early and late learning were the pair containing diseases 3 and 4, and the pair of diseases 5 and 6. In the first two blocks, disease 3 occurred three times as often as disease 4 (3:1 ratio), whereas in the last two blocks, disease 3 and 4 occurred equally often (2:2 ratio-see Table 4). The ratios were reversed for the pair containing disease 5 and 6. This gave diseases 3 and 4 unequal frequencies early and equal frequencies late and gave diseases 5 and 6 equal frequencies early and unequal frequencies late. Medin & Bettger (1991) found a positive base-rate effect for diseases 5 and 6. That is, they found more common disease responses to the PC+PR test item. However, they found an inverse base-rate effect for diseases 3 and 4 and also for diseases 7 and 8, which had been presented with a frequency ratio of 3:1 in the first two blocks and 1:3 in the last two blocks (see Table 5).[4] On this basis, Medin & Bettger (1991) argued, "disease categories were favored if they were infrequent early in training or frequent late in training, and early infrequency seems to outweigh late frequency."

One important thing to note is that Medin & Bettger (1991, see also Kruschke, 1992) suggest that participants may use base-rate information correctly. Correct use of base-rate information occurs only for the 5,6 disease pair, they say, because "...after

[4] It is not clear why there was only a slight inverse base-rate effect for diseases 1 and 2 but in some cases other researchers have also failed to find an effect (see Chapter 3-Experiment 1).

learning is complete, experience-derived base-rate information is used in the normatively correct manner." In this case, Medin & Bettger (1991) suppose that participants have finished learning the features of a category by the time the frequency shift takes place, and they then appear to be able to use frequency information properly.

Table 4. Disease structure for Medin & Bettger (1991) Experiment 2

Disease	Symptoms	Block 1 Frequency	Block 2 Frequency	Block 3 Frequency	Block 4 Frequency
1	I1 PC1	3	3	3	3
2	I1 PR1	1	1	1	1
3	I2 PC2	3	3	2	2
4	I2 PR2	1	1	2	2
5	I3 PC3	2	2	3	3
6	I3 PR3	2	2	1	1
7	I4 PC4	3	3	1	1
8	I4 PR4	1	1	3	3

Note: for 7 and 8, in which the overall relative frequencies are equal for the disease pairs, common and rare are based on the frequency ratios for the first two blocks of training.

Table 5. Response proportions to test items from Medin & Bettger (1991) Experiment 2

Disease	I	PC+PR	I+PC+PR
1	.72	.42	.56
2	.17	.47	.33
3	.58	.31	.47
4	.33	.56	.36
5	53	.58	.67
6	.36	.36	.33
7	50	.44	.53
8	.31	.56	.47

Note: proportions do not add to one because error responses (i.e., responses to a different disease pair) are excluded

Kruschke (1996) recently presented a model that proposed to account for the inverse base-rate effect (see Chapter 2). This model had two central principles: 1) when learning about categories presented at different frequencies, the more frequent category is learned first and that the rare category is learned in contrast to the more frequent category and 2) people have a bias to choose the more frequent category in classification tasks, but, "...when the bias is applied to differently encoded categories, it can be obscured and appear to be inconsistent." Kruschke (1996) was able to replicate the inverse base-rate effect in his Experiment 1 using a simplified version of the Medin and Edelson (1988) design that utilized only four diseases and six symptoms and he was able to show that a model that incorporated his two main principles was able to account for his results. However, Kruschke (1996) also noted that his studies did not provide direct evidence that the rare category is learned in terms of its distinctive features.

Questions that Remain About the Inverse Base-rate Effect

In the remainder of this thesis, after a brief discussion of several category learning models that account for the inverse base-rate effect with varying degrees of success, we describe six experiments conducted to explore various aspects of the inverse base-rate effect. The first two experiments were replications that were done to see if the inverse base-rate effect could be obtained using a standard experimental design. Experiment 1 was a replication of Kruschke's (1996) Experiment 1 and Experiment 2 was a simplified

version of the Medin & Bettger's (1991) Experiment 2. The remaining four experiments were designed to clarify the results of the preliminary work.

Although we think the inverse base-rate effect might prove to be useful in discriminating among models of category learning, we must first achieve a better understanding of the conditions under which the effect occurs and the mechanisms that are responsible for it. In Experiments 3, we inserted blocks of test trials at different points during learning to see how the inverse base-rate effect developed over time. Experiment 4 contained only a single block of test trials and served as a control to determine whether the interpolated test blocks in Experiment 3 disrupted learning. The first question we wished to address was how unequal and equal frequencies affect participants at different points during learning. Must unequal frequencies be present early in learning for it to develop, or can unequal frequencies that occur later also cause the inverse base-rate effect? Another important factor that must be considered is when is learning "complete". Do early and late unequal frequencies affect early learners differently than they affect late learners? One hypothesis we considered is that if learning is complete by the end of the first two blocks, the frequencies in early blocks should determine the responses to test items for early learners. An alternative hypothesis is that early learners may free up (attentional) resources that can be allocated to the later disease frequencies, in which case frequencies over the second two blocks should influence the responses. It might, therefore, be expected that participants who continue to make errors during the final two blocks (late learners) might show a different pattern of responses than early learners.

Next, in Experiments 5 and 6 we attempted to find support for Kruschke's (1996) proposal that the inverse base-rate effect occurs because participants first learn the common category, then learn the distinctive features of the rare category. This was done by asking participants to estimate the appropriate conditional probabilities for symptoms given diseases and diseases given symptoms after each block of learning trials. We hoped this would provide evidence as to the associations made by participants between the symptoms and diseases. We also examined how test-trial performance changed over time for both early and late learners by inserting test blocks at various points during learning, as in Experiment 3. Finally, we summarize the results of all six experiments and discuss future directions for research on the inverse base-rate effect.

CHAPTER 2

MODELING THE EFFECTS OF FREQUENCY CHANGES ON CATEGORY LEARNING

In this chapter we briefly describe models that have been proposed to account for the inverse base-rate effect. Although we did not actually perform any of the modeling reported in this thesis, the issues that the models raise are relevant to our research. One thing that makes the inverse base-rate effect so interesting is that it cannot easily be accounted for by several major memory and categorization models (Medin & Edelson, 1988; Shanks, 1992; Gluck, 1992; Kruschke, 1996).

Exemplar Models

The Context Model

Medin & Schaffer's (1978) context model of classification has been a very influential exemplar model. According to the context model, after each learning trial, the features of the stimulus and its category label are stored in a memory trace. Categorization of a new stimulus is based on its similarity to the stored traces associated with each category. Similarity to any stored trace is computed on the basis of a multiplicative similarity rule:

$$Sjk=\prod_{i=1}^{I} s_{ijk} \qquad [2]$$

where Sjk is the end product of comparing the probe with the kth exemplar of the jth category. First, the similarity of the i^{th} feature is obtained, such that if there is a perfect match $s_{ijk}=1$, and if there is not a perfect match, s_{ijk} can range between 0 and 1. Then the product of the similarities of the I features is obtained. Next the similarity of the probe to the category is obtained by summing over the k exemplars in the category. Once the similarity calculations have been made, category assignment is made using a probabilistic decision rule based on a form of Luce's (1963) choice rule:

$$P(j)= \frac{\sum_{k} Sjk}{\sum_{j=1}^{m} \sum_{k} Sjk} \qquad [3]$$

which uses the summed similarity of the probe to all the traces, s_{jk}, in a particular category as compared to the similarity of the probe to all of the stored traces in each of m categories. Using the decision rule of the context model, it can be shown that the model predicts P(common disease | PC + PR) = P(common disease | I + PC + PR) = .75. Because the model calls for the same responses for both the triplet and the conflicting pair, and predicts a positive base-rate effect for both test items, it cannot account for Medin & Edelson's (1988) inverse base-rate effect.

The Modified Context Model

In order to account for the inverse base-rate effect, Medin & Edelson (1988) proposed a modified context model that incorporated two new principles: (1) competition and (2) retrieval failure from context change. In the modified model, features or symptoms compete with each other for available strength based on their salience in learning. More salient features receive increased attention in what they refer to as representational units. As an example, when participants learn the I + PR pair, the perfect rare symptom competes with the imperfect symptom for strength in connections to the appropriate disease assignment. Errors in classification occur when the participants make predictions using the imperfect predictor. Because the perfect rare symptom is more salient for predictive purposes (i.e., using it for predictions does not result in classification errors) it receives increased attention, as compared to the imperfect predictor, in its representational unit in memory. Competition does not occur in the same way for the I + PC pair, and so it is stored as a unit in memory with fairly equal salience for each of the symptoms, along with its disease assignment. During the test phase, when a participant receives the PC + PR pair, there would be no retrieval failure due to context change for the disease associated with the rare symptom, because it is much more salient in its representational unit (e.g., I+**PR**). However, because the common symptom is only slightly more salient than the imperfect predictor in its representational unit (e.g., I+PC), there should be interference during retrieval due to the change in context from the test

item to the stored representation. Consequently, participants are more likely to predict the rare disease for the conflicting pair, because it provides the closer match to the test item.

Although Medin & Edelson (1988) gave an interesting qualitative account of their modified context model and presented some similarity values based on the differential salience of symptoms that appeared to be able to handle their findings, they never described their model in enough detail to evaluate it fully. The similarity values are never formally derived and the only justification given for their selection was that they seemed to be relationally appropriate. Currently, there is no way for the extant context model to handle the results and, until such a time when the revised model is formalized, it cannot be considered to be a viable model for handling the inverse base-rate effect.

Concept Learning by Exemplar Memorization

Myers, Lohmeier and Well (1994), using a revised version of CLEM (Chumbley, 1986), proposed a way for an exemplar model to account for the inverse base-rate effect. CLEM utilizes an error-driven mechanism, where exemplars are only stored in two situations following a categorization trial: (1) when the model's response was incorrect and (2) when there is no bias for any response, in which case the model is forced to "guess" using p=.5. Also, CLEM does not always store complete exemplars. Features of any exemplars are stored with a certain probability *r*, so the model can store fragments of exemplars. Myers et al. (1994) also discussed an alternative version of CLEM called CLEMA, in which features of an exemplar are stored on each trial with probability *r*,

even when the model responds correctly to a probe. Because this is the only difference between CLEM and CLEMA, it should be assumed that anything else pertaining to CLEM also applies to CLEMA. CLEM also allows for forgetting. When a test item activates any stored exemplar, the features of the activated traces have the probability *f* of being forgotten. Decisions by CLEM are made in an all-or-none fashion using what is called the "loudest shout rule", where evidence for a response is based on the number of traces activated when a probe is presented to the model, so the model's response is the category with the greatest number of active traces.

Myers et al.(1994) showed CLEM, in its simplest form, could not account for the inverse base-rate effect. The traces that could be potentially stored in memory are $< I, PC, \_ , D_1 >, < I, \_ , \_ , D_1 >, < \_, PC, \_ , D_1 >, < I, \_ , PR, D_2 >, < I, \_ , \_ , D_2 >,$ $< \_, \_ , PR, D_2 >$.[5] Given only the above traces, there will be an equal number of traces activated when the PC+PR pair is presented, so the evidence for D_1 or D_2 is equal and CLEM is forced to guess with p=.5. When *f* takes on a value other than 0, there are more opportunities for forgetting to occur in the stored traces. Because forgetting only occurs when stored traces are activated and traces are only activated when a matching trace is presented to the model, and because the I+PC pair is presented more frequently during learning, traces with PC are more likely to be forgotten, and the model may respond to the test item PC+PR on the basis of more traces with PR associated with the rare disease. However, in this case the model also predicts inverse responding to the triplet I+PC+PR. Myers et al. (1994) proposed a modification of CLEM. In the modified model, encoding

not only takes place when there is no evidence or equal evidence, but also takes place when the evidence for a response is weak. The probability of storing a feature when the model responds correctly then becomes:

$$r_c = r \times e^{-\alpha_{kmax}} \qquad [4]$$

where k_1 is the evidence for disease 1 and k_2 is the evidence for disease 2 and k_{max} is the larger of k_1 and k_2. Myers et al. (1994) noted that if $\alpha=0$ then the new model is equivalent to CLEMA and as α increases the model is much like CLEM. With the revised model they were able to show the appropriate response patterns for I, PC+PR, and I+PC+PR.

Connectionist Models

The Adaptive Network Model

Gluck & Bower (1988a and 1988b) were the first to try to account for the inverse base-rate effect using a simple connectionist network model. Their model, called at various times the Adaptive Network or Component Cue Model, learns associations through the changing of weights connecting a set of input nodes to a set of output nodes

[5] It should be noted that it is not necessary for CLEM to store all of the possible traces, as CLEM can respond correctly with only a subset of the traces present.

(see Apppendix D, Figure 1). The rule for spreading of activation from input to output nodes starts with the activation, a_i, of the input nodes. Activation is multiplied by the weights on the connections from the input nodes to the to output nodes (w_{ij}), and is then summed across all of the n nodes so:

$$O_j = \sum_{i=1}^{n} w_{ij} a_i \qquad [5]$$

The output of the network is then compared to the desired response (d_j), and the weights are adjusted according to the least means squared (LMS) rule or delta rule (Rescorla & Wagner, 1972; Widrow & Hoff, 1960):

$$\Delta w_{ij} = \beta(d_j - O_j)a_i \qquad [6]$$

where β is a learning rate parameter. Gluck & Bower (1988b) and Medin & Edelson (1988) asserted that because the model uses the LMS as a competitive learning mechanism where the symptoms compete for the connections to a disease it should be able to account for the inverse base-rate effect. Their claim was because the I + PC symptom pair was presented more frequently with the common disease than I+PR was with the rare disease, I should be associated more strongly with the common than the rare disease (see Table 2). Because I is more associated with the common disease (where I must compete with PC for association weight), PR should have a stronger association with the rare disease than PC has with the common disease. So, when the PC + PR pair

is presented during test, the response competition should be won by PR. Medin & Edelson (1988) pointed out that the LMS rule would not lead to this kind of responding, but instead, because I is an imperfect predictor, the model eventually learns to "ignore" I and assigns all of the weight to PC and PR. Gluck (1992) ran a simulation and demonstrated that the model, when presented with PC+PR, incorrectly predicted at asymptote an equal number of responses to the rare and common disease. Gluck (1992), furthermore, was able to show that the model, prior to asymptote, is also in error, because it consistently predicts the common disease is more likely to be chosen than the rare disease.

The Modified Adaptive Network Model

Markman (1989) also found problems with the adaptive network model and its account of the inverse base-rate effect. He claimed that the problem with the model lay in its representation of missing features. In Gluck & Bower's (1988a and 1988b) initial proposal, missing features were coded as 0's and present features were coded as 1's. Markman (1989) argued that with each feature or symptom having a unique connection to each output node, the weight of the imperfect feature to the output node was the sum of the weight of the two perfect features:

$$wI_1 = wPC_1 + wPR_1 \quad [7]$$

$$wI_2 = wPC_2 + wPR_2 \quad [8]$$

where the weights are subscripted by their connections to disease 1 and 2. Given that the network model's response to the imperfect predictor (I) is consistently the common disease, and given that the weights are initially set to 0 and cannot be less than 0, it can be shown:

$$wI_1 > wI_2 \qquad [9]$$

Substituting in from equations 7 and 8 we get the following equation:

$$wPC_1 + wPR_1 > wPC_2 + wPR_2 \qquad [10]$$

The implication of the proof is that if participants respond as the network predicts, then they should make a response of the common disease when given the conflicting pair. Therefore, the network model as presented by Gluck & Bower (1988a and 1988b) can never predict the inverse base-rate effect. Markman (1989) claimed when the model uses –1, not 0, to encode the absence of a feature, the above inequalities do not hold and the adaptive network model should predict the inverse base-rate effect. One caveat, as Kruschke (1996) notes, is that although Markman's (1989) adjusted model makes the

proper predictions for the PC +PR pair, at least pre-asymptotically, it cannot predict people's responses to the PC+PR*[6] pair.

The Distributed Stimulus Sampling Model

Gluck (1992) claimed that the problem with the original adaptive network model was that the representation of stimuli was local rather than distributed. He predicted that the model could handle the inverse base-rate effect if stimuli were represented as groups of features connected to a single output node, and when responses were based on a pattern of activations across a set of connections that added a stochastic or random element to the model (see Appendix D, Figure 2). The new variation of the adaptive network model, called the distributed stimulus sampling network model, was based on the stimulus sampling theory of Estes (1959). The proportion of connections to the output nodes after a given input is given by:

$$P_{n+1} = (1-\theta)p_n + \theta\pi \qquad [11]$$

where θ is the probability any one random element in a pool of nodes is sampled, p_n is the probability any one of the random elements in the collection of input nodes is connected to a response, and π is the probability of an outcome. The predicted change in p over n trials can be written:

[6] A symptom with a * indicates that a comparison is being made across disease pairs that do not share a

$$\Delta p=\theta(\pi-p_n) \qquad [12]$$

The model, using three groupings of input nodes to represent I, PC, and PR, showed an inverse pattern of responding to the PC + PR pair after a number of trials. One remaining problem was that when the model was able to predict the inverse base-rate effect it also incorrectly predicted the response to the I + PC + PR pair would be the rare disease.[7] Gluck (1992) was able to make adjustments to the model using a parameter he called θ' , which represented the probability of sampling elements in nonrelevant sets of input nodes and gave the model a mechanism to allow for confusability in node sampling. He also allowed for some overlap in the pools of nodes representing each symptom. The ratio of θ' to θ or the confusability of relevant and nonrelevant features, determined the behavior of the adjusted model, and Gluck (1992) was able to get the model to predict the correct pattern of responding. As Kruschke (1996) pointed out, even with the corrected distributed stimulus sampling network model using the appropriate θ ratios, Gluck (1992) still had difficulties getting the model to show a strong inverse effect for the PC+PR pair. He also offered no account of what happens in the case of the PC + PR* pair. Shanks (1992) suggested the use of an output function to attenuate the difference in the response of the network to the PC+PR pair, but so far this has not been done.

---

common imperfect predictor.

[7] The model also predicts early responses of the common disease for both I + PC + PR and PC + PR. See data in Experiment 4.

The Attentional Connectionist Model

Shanks (1992) proposed a different connectionist model to account for the inverse base-rate effect. In his model, called the Attentional Connectionist Model (ACM), Shanks (1992) stressed the role of attention to surprising elements of a stimulus. Surprise is related to the context in which the elements occur. For example, because I occurs more frequently it does not elicit any surprise when it appears in the context of the rare disease, and so attention is directed to the co-occurring element PR, which elicits surprise because of its rarity. When I +PC are presented, both elements do not elicit surprise. Shanks (1992) uses a modified delta or LMS rule, initially conceived by Wagner, (1978) to change weights in the network (see equation 4 for the original LMS rule):

$$\Delta w_{ij} = a_i(1-e_i)\beta(d_j-O_j) \qquad [13]$$

It can be seen that the LMS rule has been adjusted using $(1-e_i)$. The e_i is changed over the context in which an element of a stimulus is found and is an index of the expectedness of an element. Over time e_i changes as given by the formula:

$$\Delta e_i = \gamma_i(t_i-e_i) \qquad [14]$$

where γ_i is a free parameter of the model and t_i is a teacher that indicates to the model whether the element i is present or not. By changing over time, the formula is able to measure how expectations for stimuli change over the course of learning presentations. Shanks (1992) demonstrated that the ACM was able to produce the inverse base-rate effect for the PC+PR pair. The ACM also showed appropriate responding to the I and the I+PC+PR test items because I and PC are not surprising elements and thus, their weights are not affected by Δe_i. PR, however, is surprising, and thus its weights continue to change across the experiment. As a result, PR is more strongly connected to the rare disease than PC is to the common disease, or, as Shanks (1992) stated, the model is "...altering the readiness with which a stimulus will enter into associative relationships. ". Although initially, the revised version of the adaptive network model appeared promising, Kruschke (1996) was able to show that the ACM was now not able to handle base-rate neglect.

The Attention to Distinctive InpuT Model

Kruschke (1996) recently proposed a new connectionist model called ADIT, which is interesting because it is one of the few connectionist models that tries to incorporate relevant psychological principles. The impetus for Kruschke's (1996) ADIT was the failure of other connectionist models, including Kruschke's model ALCOVE (Kruschke, 1992), to account for the body of results in the category learning literature. Specifically, there are no current connectionist models able to handle both base-rate

neglect and the inverse base-rate effect (Kruschke, 1996; Lewandowsky, 1995). One of Kruschke's (1996) main contentions was that both effects are due to the same underlying psychological factors, so models should be able to account for them simultaneously.

In ADIT, there are two main underlying principles that constrain the use of base-rate information. One is that people use base-rate information systematically and so they have an inclination to favor the more frequent category. As Kruschke (1996) contends, "People do underemphasize the base rates relative to normative, Bayesian prescriptions, but the underweighted knowledge is consistently applied to all cases." His second proposal is that the more frequent category is learned first, giving it the distinction of being the "typical" category. The less frequent category is, therefore, learned in contrast to the typical category. In the case of the inverse base-rate effect, because both categories have the shared member I, what is attended to in the infrequent category is what makes it distinctive from the frequent category, which is PR. Kruschke's (1996) principles are based upon his own experimental work on both base-rate neglect (Kruschke, 1992) and the inverse base-rate effect (see Experiment 1).

The ADIT model is based on Gluck & Bower's (1988a and 1988b) adaptive network model with adjustments that incorporate Kruschke's (1996) two main principles.[8] First, the model contains an attentional mechanism, whereby the output for node *k* is determined by:

$$a_k^{out} = \sum_i w_{ki} \alpha_i a_i^{in} \qquad [15]$$

[8] See Kruschke and Erikson (1995)- technical report for an expansion from 2 to 5 principles.

The difference from the adaptive network model lies in the parameter α, which represents attentional strength, which can be normalized by:

$$\alpha_i = N^{-1/\eta} \qquad [16]$$

with η used as a free parameter, taking on values greater than 0. Large values of η cause attention to be strong for all featural elements, whereas small values of η produce decreasing attention to individual features when there are many features present. An interesting element in Kruschke's (1996) model is attentional strengths are adjusted prior to weight adjustments, and both types of adjustments are made based on the amount of error there is in the model's predictions; where error is defined by:

$$E = .5\Sigma(t_k - a_k^{out})^2 \qquad [17]$$

Here, t_k is a teaching signal that takes on a value of 1 if response *k* is correct and 0 if incorrect. This signal "teaches" the network to minimize error in its output. If attentional

strengths are adjusted first, weight changes are only made on the features the network finds to be relevant, as indexed by attentional strength. The model allows attention to change over time where:

$$\Delta\alpha_i=\lambda_\alpha[\Sigma(t_k-a_k^{out})w_{ki}]a_i^{in} \quad [18]$$

In this case λ_α is the attentional shift rate, which is also a free parameter of Kruschke's (1996) model. The formula allows the model to divide attention among relevant features based on what the model has learned about the category through corrective feedback, and to ignore the elements found by the model to be misleading. For example, with the rare disease, the I is misleading and so attention would be shifted to PR. Weights are then readjusted using:

$$\Delta w_{ki}= -\lambda_w(t_k-a_k^{out})\alpha_i a_i^{in} \quad [19]$$

In the formula, λ_w is another estimated parameter of the mode called the weight learning rate. Now the model is able to choose a category where the decision rule for the probability of choosing category x, using ϕ as a scaling constant is:

$$p_x=\exp(\phi a_x^{out})/\Sigma\exp(\phi a_k^{out}) \quad [20]$$

(Luce, 1963; Kruschke, 1992). Choice is based on the learning that has taken place in the network using the above attentional and weight adjustments. However, because Kruschke (1996) believes people learn something about the underlying base rates of the category structures, he provides a mechanism by which the model is able to mix the above choice probability with base-rate learning. The way the probabilities are mixed is critically contingent on the number of features, N, in the stimuli. As N increases, the use of the relative frequency information decreases in the mixed or "biased" choice probability as formalized in the equation:

$$p_x'=p_x\, b_x^{ß/(ß+N)}/\Sigma p_k b_k^{ß/(ß+N)} \quad [21]$$

where b_k reflects an accurate count of the base rates and ß is a free parameter called the base rate bias. Although Kruschke's model accounts for the data for both the inverse base-rate effect and base-rate neglect (Kruschke, 1996), the question remains as to whether the psychological principles driving the model are the actual principles operating as people learn the categories. However, it should be noted once again that modeling issues will be dealt with in future work and not in this thesis.

CHAPTER 3

PRELIMINARY RESEARCH-EXPERIMENTS 1 AND 2

Experiment 1

Experiment 1 was a replication of Kruschke's (1996) Experiment 1. We considered it important to demonstrate a robust inverse base-rate effect using a standard experimental design because attempts to demonstrate the effect have not always been successful (e.g., Medin & Edelson, 1988; Nelson, 1996; Shanks, 1992). For example, Shanks (1992) employed one design in which three disease pairs were presented for learning, two pairs which were presented at unequal frequencies and a control pair which was presented at equal frequencies. Although two disease pairs presented at unequal frequencies had the same structure (I+PC/I+PR) as employed in previous work (see Table 6), the ratio of common to rare diseases was 7:1 for one pair and 6:2 for the other. Shanks (1992) found an inverse base-rate effect for a disease pair with a 7:1 ratio, but not for a pair with a 6:2 ratio. Medin & Edelson (1988) in Experiment 2 also failed to find the inverse base-rate effect using a slightly different design. In this experiment each symptom pattern consisted of three symptoms. For one disease pair, the patterns shared only one symptom (I+PC1+PC2/I+PR1+PR2 in Experiment 2) and for a second, the patterns shared two symptoms (I1+I2+PC/I1+I2+PR). A third, control pair had no shared symptoms. The inverse base-rate effect was only found for the pair with two imperfect

predictors. Given these failures, it was necessary to demonstrate the effect before employing any manipulations to further investigate the inverse base-rate effect.

Table 6. Disease and symptom structure for Experiment 1

Disease	Symptoms	Frequency
1	I1 + PC1	3
2	I1 + PR1	1
3	I2 + PC2	3
4	I2 + PR2	1

Method

Participants. Thirty-six University of Massachusetts undergraduates took part in the experiment. They received course credit for their participation.

Stimuli and Procedure. Symptoms were presented on a computer screen, and participants were instructed to learn the correct disease assignment for each symptom pair. Six symptoms, "earaches", "skinrash", "back pain", "dizziness", "stuffy nose", and "sore muscles", were counterbalanced using the formal structure of Table 6. Four disease assignments, f, g, h, and j, were also counterbalanced across participants. In the first phase of the experiment, participants viewed pairs of symptoms and were asked to learn the disease associated with each pair. Feedback on the correct diagnosis was given on each trial on the computer screen. There were a total of 120 learning trials presented as

15 cycles of 8 trials (see Table 7 for an example trial). Presentation of the learning trials and feedback was self-paced. After responding to a learning trial participants received feedback that indicated whether or not their response was correct and then were presented with the correct response. When they finished studying the correct answer, participants pressed the spacebar to receive the next item.

Table 7. Learning trial for Experiment 1

Example trial:	DISEASE?????
	symptoms: EARACHE SORE MUSCLES
	Diagnose as one of F, G, H, J.
Feedback	WRONG!/CORRECT!
	This patient has disease ____.

After the learning phase of the experiment, participants received a test phase in which they were presented with 18 combinations of symptoms not seen during the learning phase. Each test item was presented twice, for a total of 36 test trials. Test items were presented one at a time in random order, and participants were asked to use the information from the learning phase to make a diagnosis. The test items included five standard test items I, PC, PR, PC+PR, and I+PC+PR (see Table 8) and also included four test items included by Kruschke (1996) that went across disease pairs (I+PC*, I+PR*, PC+PR*, and I+PC+PR*). Krushcke (1996) included these additional test items for

modeling purposes and we included them to provide an exact replication of his study Participants typed their responses using the keyboard and the letters f, g, h, and j were presented on the computer screen along with the test item to remind participants of the possible disease assignments. No feedback was given during the test trials.

Table 8. Test items for Experiment 1

Test Items
I
PC
PR
PC+PR
I+PC+PR
I+PC*
I+PR*
PC+PR*
I+PC+PR*

Note: C is the common disease, R is the rare disease, C* is the other common disease and R* is the other rare disease

Results and Discussion

No one was excluded from the analysis in Experiment 1 as participants had little difficulty with the task. Learning data showed that in the third 5-cycle block of learning trials participants made the correct disease assignment 94% of the time (see Table 9). The test data also provided evidence, along with the learning data, that participants had learned the disease-symptoms relationship well by the end of the learning trials and had

transferred their knowledge to the test phase. For the test items, the proportion of correct responses to perfect predictors exceeded .90 (at a proportion of .91 for the perfect common symptom and .92 for the perfect rare symptom-see Table 10).

Table 9. Proportion correct per block in Experiment 1

Block 1	Block 2	Block 3
.74	.90	.94

The test data obtained in Experiment 1 were very similar to those obtained by Kruschke (1996) Experiment 1 (see Tables 10 and 11 respectively). There was a strong inverse base-rate effect for the PC+PR test pair, ($\chi^2(1)=22.65$, $p<.001$).[9] The imperfect predictor was classified 78% of the time as the common disease and only 11% of the time as the rare disease. We were also able to replicate the inverse base-rate effect for the PC+PR* pair ($\chi^2(1)=16.13$, $p<.001$) which is also consistent with Kruschke's (1996) findings. Although Kruschke (1996) and others (Medin & Bettger, 1991; Medin & Edelson, 1988) found a positive base-rate effect for the I+PC+PR triplet, this was not significant in Experiment 1 ($\chi^2(1)=.71$, $p>.05$). However, responses to the I+PC+PR* pair showed a positive base-rate bias ($\chi^2(1)=11.88$, $p<.005$), again consistent with

[9] We used χ^2 to test the significance of the effect as Kruschke (1996) did throughout his experiments. However, Kruschke (1996) notes that χ^2 is not truly the appropriate measure, because "the data include repeated measures from the same participants...the data violate the assumption of independence..." Wickens (1989) presented a measure for correcting for such violations, but Kruschke (1996) found that with such large χ^2 values the correction was not critical.

Table 10. Response proportions obtained in Experiment 1

Test Items	C	R	C*	R*
I	.775	.110	.055	.030
PC	.910	.040	.035	.015
PR	.010	.920	.045	.020
PC+PR	.280	.680	.005	.035
I+PC+PR	.510	.450	.010	.025
I+PC*	.350	.045	.585	.020
I+PR*	.165	.080	.050	.705
PC+PR*	.310	.025	.020	.640
I+PC+PR*	.615	.035	.020	.330

Note: C is the common disease, R is the rare disease, C* is the common disease from the other disease pair and R* is the rare disease from the other disease pair

Table 11. Response proportions-Kruschke (1996)

Test Items	C	R	C*	R*
I	.746	.174	.049	.031
PC	.933	.031	.031	.004
PR	.040	.911	.018	.031
PC+PR	.353	.612	.022	.013
I+PC+PR	.580	.402	.013	.004
I+PC*	.406	.080	.469	.045
I+PR*	.219	.085	.031	.665
PC+PR*	.353	.027	.058	.563
I+PC+PR*	.719	.036	.036	.210

Kruschke's (1996) results. In summary, the major results obtained by Kruschke (1996) were also found in this experiment, and the inverse base-rate effect was demonstrated using the general design that will be employed in the remainder of the studies.

Experiment 2

After demonstrating a robust inverse base-rate effect in Experiment 1, we focused on how the inverse base-rate developed when frequencies were varied at different points during learning. Experiment 2 was a reduced version of Medin & Bettger's (1991) study (see Table 12 for disease structure). Changes were made in the Medin & Bettger (1991) design to address concerns about their study, for example, the possible confound of attention to multiple embedded problem structures with the effects of frequency, the heavy learning load placed on participants, and their failure to examine individual learning curves and to ask for early responses to test items.

To elaborate, there were a number of possible problems with the design of Medin & Bettger's (1991) Experiment 2. First, it was possible in their study that the 5,6 disease pair (see Table 4) received more attention in the first two blocks because it was the only pair presented with equal frequencies (in these blocks). It was not clear if Medin & Bettger (1991) did not find an inverse base-rate effect for this pair because of the equal frequencies in the first two blocks, or because the pair stood out from the other pairs and so received more attention. A similar problem was present for diseases 3 and 4 in the second two blocks. Because of these problems it is not easy to interpret Medin & Bettger's (1991) results. Therefore, in Experiment 2 we simply isolated the two most

critical conditions of Medin & Bettger's (1991) study (diseases 3 and 4 and diseases 5 and 6-see Table 4) so participants would attend equally to each pair and not be influenced by having to learn multiple embedded problems.

Medin & Bettger (1991) also stated , "...predictions hinge on the majority of learning taking place over the first two blocks." However, in the Medin & Bettger (1991) study, learning was still taking place in the fourth block for some category pair members for at least some participants. Moreover, because Medin & Bettger (1991) did not look for differences in test items for participants who learned quickly and those who learned slowly, claims that they made about the effects of frequencies presented early and late in learning need to be studied more carefully. In their test trials Medin & Bettger (1991) found that the proportion of correct responses to the perfect predictors of disease 3 and 4 were .81 and .71 respectively, whereas in other studies the proportion correct has been typically greater than .90 (see Table 7). Poor performance in the Medin & Bettger (1991) study might have been due to the heavy learning load placed on participants. These issues provided us with a reason to present only the critical two pairs to participants to reduce the learning load and to look directly at participants' learning curves over time by examining data from both "early" and "late" learners. This enabled us to investigate what happened when learning took place prior to the frequency shift as opposed to what happened when learning continued after the frequency shift. As stated previously, learning before the shift might have caused the early frequencies to have the greatest effect on responses to the test items. The other plausible alternative was that early learners would have more resources available in the second half of learning, and they might have paid more attention to these disease frequencies. Therefore, it was possible

that these later frequencies might have more strongly influenced early learners. It was also unclear as to which frequencies (early or late) should influence later learners, although it seems more likely that late frequencies would more strongly influence their performance. It is therefore highly probable that early and late learners have different response patterns.

Method

Participants. Forty-one University of Massachusetts undergraduates took part in the experiment and received course credit for their participation.

Stimuli and Procedure. Learning trials and test trials were the same as those used in Experiment 1. The only changes made from Experiment 1 in the learning trials were to the disease frequencies (see Table 12) and to the number of learning trials which were increased from 120 to 128; that is, four blocks each consisting of four eight-trial cycles.

Table 12. Disease structure for Experiment 2

Disease	Symptoms	Block 1 Frequency	Block 2 Frequency	Block 3 Frequency	Block 4 Frequency
C1	I1 PC1	3	3	2	2
R1	I1 PR1	1	1	2	2
C2	I2 PC2	2	2	3	3
R2	I2 PR2	2	2	1	1

Note: 1 is used as an index for the disease pair initially presented at unequal frequencies and 2 is used as an index for the disease pair initially presented at equal frequencies.

Following the learning phase, participants were given a test phase in which they were presented with the same 18 test items used in Experiment 1 (see Table 8). As in Experiment 1, each test item was presented twice, for a total of 36 test trials.

Results and Discussion

Five participants were dropped from the analyses because the did not reach a criterion of 100% correct responses in at least one cycle of learning trials (see Table 13). Of the remaining thirty-six participants, 17 were classified as early learners, 7 were classified as late learners and 12 were excluded from the early/late analysis because they reached criterion during the middle of the learning trials (see Table 13). Early learners were defined as those whose learning appeared to be complete by the end of the first block of 32 trials.[10] No more than one error per cycle could be made after the first block of 32 trials. Late learners were defined as participants who continued to make errors into the fourth block of learning trials. Early had a significantly higher proportion of correct responses than late learners for the first three blocks (separate variance $t(10.3)=2.30$, $p=.04$; $t(7.3)=7.85$, $p<.001$; $t(7.3)=5.71$, $p=.001$ for blocks 1, 2, and 3, respectively).

[10] Binder and Estes (1966) made a similar division by separating participants based on the number of errors each participants made, although other definitions are possible (see Experiments 4 and 5).

Table 13. Proportion correct per block in Experiment 2

	1	2	3	4
Overall	.68	.88	.96	.98
	(.03)	(.02)	(.01)	(.01)
Early	.76	.97	.99	.99
	(.03)	(.01)	(.01)	(0)
Late	.63	.78	.87	.96
	(.05)	(.02)	(.02)	(.02)

Note: Hereafter, numbers in parentheses are standard errors.

For the test trials, the results of the present study are somewhat inconsistent with those of Medin & Bettger (1991). Tables 14 and 15 display the results for the critical test items. In the Medin & Bettger (1991) study, participants showed an inverse base-rate effect (.31 to .56) for the pair initially presented at unequal frequencies and later presented at equal frequencies, hereafter called the "unequal early" condition, and a positive base-rate effect (.58 to .36) for the pair initially presented at equal frequencies and later presented at unequal frequencies, hereafter called the "unequal late" condition. Although not significant, our initial results in the unequal early condition are consistent with Medin & Bettger (1991), with a nonsignificant trend toward an inverse base-rate effect (.40 to .60, $p>.05$-see Table 14). However, results in the unequal late condition did not show correct use of base-rate information but instead showed nearly equal responding with response proportions of .44 to the common disease and .51 to the rare disease. This may have been due to the reduced learning load in Experiment 2. Participants in Experiment 2 might have learned the material more quickly than Medin & Bettger's (1991) participants, and so may have been responding in the unequal late condition based on the early equal frequencies, whereas Medin & Bettger's (1991) participants may have

taken longer to learn the material and so may have been responding based on the later unequal frequencies. It is also possible that Medin and Bettger's (1991) participants were influenced by the multiple embedded problem structures.

Table 14. Response proportions to test items from Medin & Bettger (1991) Experiment 2

Test Items	Disease 3	Disease 4	Disease 5	Disease 6
I	.58	.33	.53	.36
PC+PR	.31	.56	.58	.36
I+PC+PR	.47	.36	.67	.33

Table 15. Response proportions to test items for Experiment 2

Test Items	Disease 1	Disease 2	Disease 3	Disease 4
I	.81	.15	.50	.40
PC+PR	.40	.60	.44	.51
I+PC+PR	.61	.38	.56	.42

Note: the full set of response proportions can be found in Appendix A1

Because our findings were different than those of Medin & Bettger (1991) for the PC+PR pair in the unequal late condition, we examined the results for early and late learners. The initial expectation was that early learners should show a positive base-rate effect for the PC+PR item in the unequal late condition, given that for these participants, most of their learning occurs during a period of equal frequencies. According to Kruschke (1996), in this situation, participants should use base-rate information accurately.

The 17 early learners did not show this result (see Table 16). In fact, what was found was a nonsignificant trend towards an inverse base-rate effect (.41 to .59, p>.05). Surprisingly, it was the late learners who showed a nonsignificant positive base-rate trend (.43 to.36, p>.05). However, we did not draw any conclusions from these preliminary results because there were only 7 late learners and they exhibited such a high error rate (21%) in response to this test item.

The results of Experiment 2 suggested that by simplifying Medin & Bettger's (1988) study we might learn more about the inverse base-rate effect, how the effect develops over time, and how it is influenced by differing frequencies during

Table 16. Response to PC+PR for the unequal late condition

	C	R	Difference
Early learners (N=17)	.41	.59	.18
Late learners (N=7)	.43	.36	.07

learning. The results also gave an indication that some changes in the design[11] might be necessary to more closely examine the development of the inverse base-rate effect. Finally, this study showed us that if we wished to compare the results of early and late learners we would need to increase the number of participants run in subsequent experiments.

[11] These changes will be expanded upon in Chapter 4.

CHAPTER 4

EXAMINING THE DEVELOPMENT OF THE INVERSE BASE-RATE EFFECT IN EARLY AND LATE LEARNERS

Experiments 3 and 4

The purpose of Experiments 3 and 4 was to learn more about the development of the inverse base-rate effect by examining how early learners differed from late learners. As stated previously, we expected early and late learners to be influenced differently by changing disease frequencies. We also investigated how the inverse base-rate effect develops over time by probing participants at different points in learning.

Medin & Florian (1992) pointed out that although in the past there has been an emphasis on trying to explain performance on test trials, more recently there has been an increased emphasis on trying to account for performance during learning. Also, we should note that there has been other recent work that has emphasized the importance of examining response patterns during learning and that has shown that later learning trials can have an effect on the representation that is formed by the learner (Erikson & Kruschke,1998). To explore the claims of researchers who point to the differential influence of early learning, and the continued influence of learning across all training trials, in Experiment 3 test items were presented to the participants after each block of 32 learning trials to track learning over time and to trace the development of the inverse base-rate effect for both early and late learners. Experiment 4 participants were only probed at the end of all four blocks of learning trials. Because participants in Experiment

4 only received test trials after all 128 learning trials had been completed, we were able to compare performance on the final set of test trials across the two experiments to see if the intervening test trials in Experiment 3 affected performance.

Because participants could not a priori be assigned to the early and late learning groups it was deemed necessary to run a much larger group of participants. We also more precisely redefined early learners as those who made no errors for two consecutive cycles of 8 learning trials prior to the frequency shift at the end of block 2 and defined late learners as those who first made no errors for two consecutive cycles of eight learning trials after the frequency shift. This was done so early learners could be clearly contrasted with those who met this criterion when the opposite frequencies were in effect. This allowed for a more unambiguous test of the hypothesis that response patterns are a function of the underlying frequency ratios during which learning is completed.

One change in design that was made was the number of test items seen by each participant. In Experiments 1 and 2 participants were given only two presentations of each test item. It was difficult with only two responses per participant to examine individual response patterns for inverse trends. So, to examine these individual trends and specifically to look at whether individual participants show the inverse base-rate effect for the PC+PR test item and positive base-rate bias for I+PC+PR item, or if these effects were an artifact of averaging across participants, we presented each item four times in each test block.

The data collected in Experiment 3 should be helpful in testing models. For example, ADIT, because of certain inherent characteristics, has a protracted learning period. Kruschke (1996) noted, "In early training, ...the model performed much worse

than human learners; for example, in the first block of training, human learners attained 68% and 49% correct for the common and rare diseases, respectively, but the model showed only 54% and 16% correct, respectively. ...early training performance on the rare diseases might be elevated in humans because of mechanisms not implemented in the model." It takes ADIT close to four blocks to reach the point where it responds correctly on over 90% of the test trials, unlike many human participants who reach this level of performance relatively early in training. Therefore, when the model is presented with the learning trials and test trials which we use in Experiment 3, it should perform more like late learners, not early learners.

Method

Participants. Fifty University of Massachusetts undergraduates took part in Experiment 3 and 72 took part in Experiment 4. They received course credit for their participation.

Stimuli and Procedure. In these studies, participants were again asked to learn how six symptoms (earache, back pain, dizziness, skin rash, sore muscles, and stuffy nose) were paired with four diseases (f, g, h, j) using the same disease frequency structure as in Experiment 2 (see Table 11). For the first two blocks of learning trials, two diseases and three symptoms were used in the unequal base-rate condition, in which the common disease occurred three times as often as the rare disease. The other two diseases and three symptoms were paired with the equal base-rate condition. After the first two blocks of

learning trials, the diseases that had been presented in the unequal base-rate condition were presented for the final two blocks with equal frequencies and those that had been presented with equal frequencies were presented with unequal (3:1) frequencies. The presentation of learning trials was the same as in Experiment 1 (see Table 7).

Table 17. Test items for Experiments 3 and 4

Test Items
1) I
2) PC
3) PR
4) PC + PR
5) I + PC + PR
6) I+PC
7) I+PR

The seven critical test items for both Experiment 3 and Experiment 4 are shown in Table 17. They included the critical five test items from Experiments 1 and 2 that tested within disease pairs. Two new test items (I+PC and I+PR) which were the same as the learning trials, were also included to see if participants were attending to the test items during the test trials and to examine Kruschke's (1996) proposal that the common category is learned first which causes the rare category to be learned in contrast to it. Following each block of 32 learning trials in Experiment 3, participants were presented with these test items. Each test item was presented four times for both disease pairs, for a total of 56 test trials after each block or a total of 224 test trials in Experiment 3. In

Experiment 4 the 56 test trials were presented only once at the end of the 128 learning trials.

Learning Data

In Experiment 3 one participant was dropped and in Experiment 4 eleven participants were dropped for not reaching the learning criterion of two cycles of 100% correct responses by the end of the learning trials. If the intervening test trials had substantially increased the difficulty of learning the disease assignments in Experiment 3 we would have expected to see a greater number of participants dropped from Experiment 3 than from Experiment 4, but the opposite was true. There were 21 early learners and 28 late learners in Experiment 3, and 36 early learners and 25 late learners in Experiment 4. Although there appear to be more early learners in Experiment 4 than in Experiment 3 this difference was not significant ($\chi^2(1)=3.00$). As can be seen in Tables 18 and 19 for both Experiments 3 and 4 the proportion correct for late learners was much lower than that of early learners for the first two blocks, but approached the levels of the early learners in the last two blocks. There were significant type of learner by block interaction for both Experiments 3 and 4 ($F(3,141)=9.27$, $p<.001$ and $F(3,177)=16.43$, $p<.001$, respectively). Learning took place at about the same rate in both Experiments 3 and 4 both overall and for early and late learners separately, providing initial evidence that the intervening learning trials in Experiment 3 did not interfere with learning.

Table 18. Proportion correct per block in Experiment 3

	Block			
	1	2	3	4
Overall (N=49)	.63 (.02)	.92 (.01)	.96 (.01)	.97 (.01)
Early (N=21)	.72 (.02)	.97 (.01)	.98 (.01)	.98 (.01)
Late (N=28)	.56 (.03)	.89 (.02)	.95 (.01)	.96 (.01)

Table 19. Proportion correct per block in Experiment 4

	Block			
	1	2	3	4
Overall (N=61)	.66 (.02)	.90 (.01)	.96 (.01)	.98 (0)
Early (N=36)	.73 (.02)	.97 (.01)	.98 (.01)	.99 (0)
Late (N=25)	.57 (.02)	.81 (.02)	.93 (.02)	.97 (.01)

Test Data

The Inverse Base-rate Effect.[12] In the unequal early condition, the inverse base-rate effect clearly developed by the end of second block in Experiment 3 (effect size .29, t(48)= -2.73, p=.009-see Table 20) and remained at about the same size even after the disease frequencies were equal (effect size .30, t(48)= -2.88, p=.006; effect size .29, t(48)= -2.52, p=.015, for blocks 3 and 4 respectively). This was a good indication that the inverse base-rate effect developed quite early in learning. There was a large inverse base-rate effect in Experiment 4 (effect size .57, t(60)=-7.36, p<.001). There was no inverse base-rate effect in any block in the unequal late condition in Experiments 3 and 4, which was evidence that unequal frequencies, in general, needed to be present early in learning for the inverse base-rate effect to develop.

Table 20. Response proportions to test item PC+PR

		Unequal early				Unequal late		
Experiment	Block	C	R	Diff		C	R	Diff
3	1	.44	.43	.01		.45	.44	.01
	2	.35	.64	−.29	**	.50	.46	.04
	3	.35	.65	−.30	**	.48	.52	−.04
	4	.35	.64	−.29	*	.48	.51	−.03
4	4	.21	.78	−.57	***	.45	.54	−.09

Note: hereafter significance at .05 will be denoted by *, at .01 by **, and at .001 by ***.

[12] The complete results for test items I, PC, PR, PC+PR, I+PC+PR, I+PC, and I+PR, including error responses to inappropriate diseases, can be found in Appendix B Tables B1-B7.

The Inverse Base-rate Effect in Early and Late Learners. Early learners showed the inverse base-rate effect clearly by the end of the second block in Experiment 3 (effect size .50, t(20)=-3.52, p=.002-see Table 21) and the effect remained constant throughout the final two blocks in Experiment 3 (effect size .50, t(20)= -3.52, p=.002; effect size.60, t(20)= -4.11, p=.001 for blocks 3 and 4 respectively). The effect was also present for early learners in Experiment 4 (effect size .61, t(35)= -6.25, p<.001) and this was virtually identical to the effect size at the end of block 4 in Experiment 3. This suggests that early frequency patterns exert a strong influence on early learners because most of their learning takes place under those frequencies. It also indicates that these early frequencies continue to exert an influence even after the frequencies become equal in block 3, indicating that for early learners, once learning is completed, the influence of the frequency patterns under which learning took place exert the greatest influence on responses. They also indicate that the intervening test trials in Experiment 3 do not appear to influence the development of the inverse base-rate effect in early learners.

Table 21. Response proportions to test item PC+PR for early and late learners

		Early learners						Late learners					
		Unequal early			Unequal late			Unequal early			Unequal late		
Exp	Block	C	R	Diff	C	R	Diff	C	R	Diff	C	R	Diff
3	1	.42	.51	−.09	.38	.57	−.19	.46	.37	.09	.50	.34	.16
	2	.25	.75	−.50**	.45	.52	−.07	.42	.56	−.14	.53	.42	.09
	3	.25	.75	−.50**	.43	.57	−.14	.42	.57	−.15	.52	.47	.05
	4	.20	.80	−.60***	.40	.60	−.20	.46	.52	−.06	.54	.45	.09
4	4	.19	.81	−.61***	.45	.55	−.10	.24	.74	−.50***	.45	.53	−.08

The most puzzling result, however, was that late learners in Experiment 3 never showed the inverse base-rate effect, even though late learners in Experiment 4 showed a large inverse base-rate effect (effect size .50, $t(24) = -4.00$, $p = .001$). It is not clear why the intervening test trials in Experiment 3 changed the pattern of responding. As stated earlier, there was essentially no difference in the learning curves for late learners in Experiment 3 and late learners in Experiment 4, and so the difference in response patterns cannot be attributed to a simple learning difference. It might be the case that viewing the test trials resulted in some form of self-feedback which influenced the internal representation of the disease and symptom frequencies. However, it is premature at this point to draw any conclusions because, overall, the pattern of responding in Experiment 3 (for the final block of test items) was so similar to the pattern of responses in Experiment 4 and if the category representation had been altered, it should have also affected response patterns to other test items.

I+PC+PR Triplet. The I+PC+PR test triplet (see Table 22) also provided some interesting results. The general finding in the literature has been that when presented with the I+PC+PR test item, participants typically choose the more common disease, exhibiting what is thought to be correct use of base-rate information. Overall in Experiment 3 in the unequal early condition, there appeared to be correct use of base-rate information for the first two blocks, with a positive base-rate bias being present for these blocks (effect size .35, $t(48) = 3.81$, $p = .001$, effect size .23, $t(48) = 2.11$, $p < .05$). The positive base-rate bias disappeared after the third block, which corresponded to the frequency shift and the bias was also not present in Experiment 4. This indicates that

later frequencies exert an influence on response patterns. In the unequal late condition there was also a marginally significant inverse triplet base-rate effect in block 3 (t(48)= -2.02, p=.05).

Table 22. Response proportions to test item I+PC+PR

		Unequal early				Unequal late			
Experiment	Block	C	R	Diff		C	R	Diff	
3	1	.65	.30	.35	***	.34	.53	−.19	
	2	.61	.38	.23	*	.43	.56	−.13	
	3	.49	.51	−.02		.38	.61	−.23	*
	4	.48	.50	−.02		.44	.56	−.12	
4	4	.44	.54	−.10		.42	.56	−.14	

The I+PC+PR Triplet and Early and Late Learners. As can be seen in Table 23, for I+PC+PR triplet in the unequal early condition, early learners in Experiment 3 showed a non-significant positive base-rate trend in the first block (.63 to .32, t(20)=1.71, p>.05), but an inverse pattern of responding for the last two blocks (.34 to .66, t(20)= -1.68, p>.05; .29 to.71, t(20)=-2.76, p=.01). They were influenced by the frequencies throughout the course of learning. Early learners in Experiment 4 exhibited a non-significant inverse pattern similar to the final pattern of early learners in Experiment 3 (.40 to .58, t(35)=-1.72, p>.05). Late learners in Experiment 3 in the unequal early condition showed a positive base-rate bias in the first two blocks of learning (effect size .38-t(27)=3.12, p=.004; effect size .41-t(27)=3.46, p=.002) that remained positive through

blocks 3 and 4[13] indicating that, in contrast to the early learners in the unequal early condition, once their response pattern had been set early in learning it was not strongly influenced by later frequency shifts.

Table 23. Response proportions to test item I+PC+PR for early and late learners

		Early learners						Late learners					
		Unequal early			Unequal late			Unequal early			Unequal late		
Exp	Block	C	R	Diff	C	R	Diff	C	R	Dif	C	R	Diff
3	1	.64	.32	.32	.26	.68	−.42**	.66	.28	.3 **	.40	.41	−.01
	2	.49	.50	−.01	.42	.58	−.16	.70	.29	.4 **	.44	.54	−.10
	3	.34	.66	−.32	.31	.68	−.37*	.60	.40	.2	.44	.55	−.11
	4	.29	.71	−.42**	.39	.61	−.22	.63	.35	.2	.48	.52	−.04
4	4	.40	.58	−.18	.42	.58	−.16	.51	.48	.0	.43	.54	−.11

For the I+PC+PR triplet in the unequal late condition, late learners in Experiment 3 show no bias for any block, indicating again that the early equal frequencies exert the strongest influence on late learners and that any later frequency shift appears to have little effect on these learners. Late learners in Experiment 4 also show no bias for the triplet.

However, it is important to mention that the I+PC+PR triplet for the pair of diseases in the unequal late condition highlighted a potential problem with the I+PC+PR test items. The response patterns of early learners in Experiment 3 showed an inverse pattern of responding to the probe for blocks 1 (effect size .42, $t(20) = -2.70$, $p = .01$),

[13] Block 3 and block 4 were in the same direction as the effect in blocks 1 and 2 but they were not significant (p=.17 and .05 respectively).

which was also present in block 3 (effect size .37, t(20)= -2.21, p=.04).[14] However, given the equal frequencies in the first half of training, there was no reason for participants in Experiment 3 to show any preference for one disease or another in the first two blocks. When the stimuli were examined closely, a potential problem was discovered with the counterbalancing. The I, PC, and PR symptoms were presented on the computer screen in a vertical array (see Table 6). Given that the triplet had to be presented four times per block, there was no way of presenting all 6 possible vertical orderings of three stimuli with only four presentations and, it was decided to present the imperfect predictor at the top of the list of symptoms for two trials and at the bottom for the other two. This gave the following two presentation orders, I+PR+PC/PR+PC+I. It should be noted that in these orderings the perfect rare symptom always appeared above the perfect common symptom. It is possible that participants, in responding quickly, might have scanned down to the first perfectly predictive symptom, which, given the stimuli, was always the perfect common symptom. This could account for the observed bias in responding after block 1. To correct the problem, the following orderings, PC+PR+I/ PR+PC+I, could be used to eliminate any bias due to presentation order to allow us to examine if there is truly a difference in responding for early and late learners and if early learners are influenced by later frequency shifts, or if the effects observed were simply due to improper counterbalancing. These orderings were used in Experiments 5 and 6.

The Imperfect Predictor. Participants showed a positive base-rate bias across all four blocks in Experiment 3 and after the final block in Experiment 4 for the imperfect predictor in the unequal early condition (see Table 24), indicating that correct use of base-

[14] These t values are only marginally significant and should be treated with some caution.

rate information for this disease pair developed quite early in learning. Both early and late learners in Experiments 3 and 4 showed the same pattern of responding for the imperfect predictor in this disease pair (see Table 25) although the bias was not present until the second block for late learners. For the unequal late condition in Experiment 3, the responses in the first two blocks, both overall and for early and late learners separately were also in accord with the base rates, which was expected.

Table 24. Response proportions for test item I

		Unequal early				Unequal late		
Experiment	Block	C	R	Diff		C	R	Diff
3	1	.64	.26	.40	***	.42	.40	.02
	2	.74	.20	.54	***	.52	.42	.10
	3	.75	.24	.51	***	.56	.42	.14
	4	.72	.26	.46	***	.58	.39	.19
4	4	.73	.26	.47	***	.44	.53	−.09

Table 25. Response proportions to I for early and late learners

		Early learners							Late learners						
		Unequal early				Unequal late			Unequal early				Unequal late		
Exp	Block	C	R	Diff		C	R	Diff	C	R	Diff		C	R	Diff
3	1	.79	.20	.59	***	.49	.38	.11	.54	.30	.24		.37	.44	−.07
	2	.81	.18	.63	***	.58	.41	.17	.69	.22	.47	**	.47	.44	.03
	3	.79	.20	.59	***	.60	.38	.22	.71	.26	.45	**	.53	.45	.08
	4	.80	.20	.60	***	.57	.41	.16	.66	.30	.36	*	.59	.38	.21
4	4	.76	.23	.53	***	.40	.58	−.18	.67	.31	.36	*	.50	.44	.06

The Perfect Common and Perfect Rare Predictors. Participants learned the correct disease assignment for the more common perfect predictor and for the rare perfect predictor in the unequal early condition both overall and for early and late learners by the end of block1. The responses to all of the perfect predictors were over 90% by the end of learning, and this provided additional evidence participants learned the material and transferred their knowledge from the learning trials to the test trials.

Table 26. Response proportions for test item PC

		Unequal early			Unequal late		
Experiment	Block	C	R	Diff	C	R	Diff
3	1	.66	.17	.49 ***	.58	.20	.38 ***
	2	.91	.05	.86 ***	.84	.10	.74 ***
	3	.94	.02	.92 ***	.94	.03	.91 ***
	4	.93	.05	.88 ***	.96	.03	.93 ***
4	4	.96	.04	.92 ***	.95	.05	.90 ***

Table 27. Response proportions to test item PC for early and late learners

		Early learners						Late learners					
		Unequal early			Unequal late			Unequal early			Unequal late		
Exp	Block	C	R	Dif	C	R	Diff	C	R	Diff	C	R	Diff
3	1	.81	.13	.6 ***	.68	.24	.44**	.55	.21	.34**	.51	.18	.33**
	2	.98	.02	.9 ***	.91	.08	.83***	.86	.06	.80***	.80	.12	.68***
	3	.96	.02	.9 ***	.94	.02	.92***	.93	.02	.91***	.95	.04	.91***
	4	.96	.04	.9 ***	.99	.01	.98***	.91	.05	.86***	.95	.05	.90***
4	4	.95	.05	.9 ***	.99	.01	.98***	.96	.04	.92***	.89	.11	.78***

Table 28. Response proportions to test item PR

		Unequal early				Unequal late			
Experiment	Block	C	R	Diff		C	R	Diff	
3	1	.13	.65	−.52	***	.25	.63	−.38	***
	2	.05	.87	−.82	***	.10	.86	−.76	***
	3	.02	.95	−.93	***	.03	.93	−.90	***
	4	.01	.97	−.96	***	.04	.93	−.89	***
4	4	.03	.96	−.93	***	.03	.97	−.94	***

Table 29. Response proportions to test item PR for early and late learners

		Early learners						Late learners					
		Unequal early			Unequal late			Unequal early			Unequal late		
Exp	Blk	C	R	Diff	C	R	Diff	C	R	Diff	C	R	Diff
3	1	.07	.83	−.76 ***	.14	.77	−.63 ***	.18	.5	−.34**	.32	.53	−.21
	2	0	1.00	−1.00 ***	.01	.98	−.97 ***	.09	.7	−.69***	.16	.77	−.61***
	3	.01	.98	−.97 ***	.02	.96	−.94 ***	.03	.9	−.91***	.04	.90	−.86***
	4	0	.99	−.99 ***	.01	.98	−.97 ***	.01	.9	−.95***	.05	.90	−.85***
4	4	.03	.97	−.94 ***	.01	.98	−.97 ***	.04	.9	−.90***	.04	.95	−.91***

I+PC and I+PR. Finally, there were the two test items, I+PC and I+PR (see Tables B6 and B7 in Appendix B), that were included to test Kruschke's proposal that "the primary role of base rates … is to cause the high-frequency categories to be learned before the low-frequency categories.". Did participants learn I+PC before they learned I+PR, so that I+PR would be learned in contrast to I+PC? Participants made the correct assignment to I+PC for the unequal early condition in 81% of the test trials after only one block of learning trials in Experiment 3, which was much higher than the 52% correct

responses to I+PR for the disease pair initially presented at unequal frequencies[15] ($t(48)=5.27$, $p<.001$). So, I+PR appeared to be learned after I+PC, as Kruschke (1996) asserted.

Discussion

What can be concluded so far from the results? First, the inverse base-rate effect developed in the unequal early condition early in learning (by the end of the second block in Experiment 3) and persisted once it developed. Intervening test trials did not appear to disrupt its development. The intervening test trials in Experiment 3 did not slow learning in comparison to Experiment 4. However, in contrast to the results of Medin & Bettger (1991), there was also evidence that the inverse base-rate effect only occurred if the majority of learning took place during a period of unequal frequencies, as shown by the results of early and late learners in Experiment 3. Whereas early learners showed the development of the inverse base-rate effect by the end of block 2 in the unequal early condition, late learners did not show this pattern. One caveat we should note is that, when not interrupted by test trials, late learners do show an inverse base-rate effect. Although a logical explanation would be that learning is more disrupted and therefore more drawn out for late learners when intervening test trials were given and that in this case late learners would provide a better or more distinct contrast to early learners than

[15] It should be noted that we are only comparing the diseases initially presented at unequal frequencies because there is no theoretical reason why the diseases that were equally frequent would be learned at different rates.

the late learners in Experiment 4, this did not appear to be the case when the learning data were examined.

For the triplet results, there was evidence that later frequencies can have an influence, even once learning has been completed. Also, if there really does prove to be an interaction of frequency with responses for the early and late learners for the I+PC+PR test item in further research, it would provide additional evidence that frequencies affect early and late learners differently. However, this effect must be established in studies in which the possible confound of display position has been removed. This confound was removed in Experiments 5 and 6 to further explore this finding.

Finally for the remaining test item results, there is support from Experiment 3 for Kruschke's hypothesis that the more frequent category is learned first (I+PC versus I+PR). It also appears, because the imperfect predictor is more closely associated with the more frequent category and is not strongly associated with the rare category in both Experiments 3 and 4, the rare disease must be learned mostly by its distinguishing feature PR. If the rare disease was learned in terms of both of its features, I and PR, there should have been a greater number of rare disease responses to the imperfect predictor. This seems to support Kruschke's (1996) claim that what is learned about the less frequent category is what distinguishes it from the frequent category. The imperfect predictor also provides more evidence that later frequencies influence participants after learning is completed.

There were several methodological issues raised by these experiments. Although there were a few differences in the final response patterns from Experiment 3 and 4, it appeared as though the intervening test trials did not interfere greatly with learning as can

be seen in the learning data and in the number of dropped participants. Even though there were slight differences when intervening test trials were present, we should not forget the value that they add to our research. Because the intervening test trials allowed us to track response patterns across time, they are a valuable addition to the previous work of Medin & Bettger (1991), Kruschke(1992), and others. They also enabled us to see how early and late learners' response patterns diverged at different points in learning. Because of these benefits, we continued to employ this method in both Experiments 5 and 6.

Lastly, it should be noted that there are some possible objections to using the distinction of early and late learners. It is possible to argue that we were simply looking at smart participants or those who make an effort versus those who made little effort to learn the material. Because of our learning criterion, all participants in the analysis performed at over 90% accuracy by the end of the experiments and so they did not include participants who could not or refused to learn the material and who were dropped from the analyses. Therefore, we felt that because of the benefits this analysis provides, namely, to see how underlying frequencies effect categorization patterns under which learning takes place, that it was a valid distinction to make and that we should continue to use this method for dividing participant in future analyses performed in Experiments 5 and 6.

CHAPTER 5

CONDITIONAL PROBABILITY ESTIMATES AS A MEASURE OF WHAT IS KNOWN ABOUT CATEGORY FEATURES

Experiments 5 and 6

The primary purpose of Experiments 5 and 6 was to provide a direct test of what participants can report about features given a specific disease and what they can report about diseases given a specific feature. One of Kruschke's two main principles is the assertion that participants learn the characteristics of the rare category in contrast to those of the common category. That is, they learn the characteristics of the common category first and then learn the feature(s) of the rare category that distinguishes it from the common category. However, Kruschke (1996) stated that his "experiments...provided only indirect evidence that the rare categories are encoded primarily by their distinctive features...". Without direct evidence to support this hypothesis, it is difficult to evaluate Kruschke's ADIT model. In contrast, Myers et al. (1994) proposed that the inverse base-rate effect is due to the greater frequency with which $< \_PC, \_, D_1 >$ traces are activated, because this activation causes them to be more susceptible to forgetting. These two contradictory hypotheses raise the question as to what factor(s) underlie the inverse base-rate effect. Are rare categories encoded primarily by their distinctive features, as Kruschke (1996) proposes, or, as Myers et al.(1994) assert, are features initially encoded for both the rare and common diseases, with traces only being lost because of the more

frequent activation of common category traces? Experiments 5 and 6 attempt to address the question of exactly what features are encoded for rare and common diseases, and how frequencies and frequency changes influence the features that are encoded at various points in learning.

In order to examine what participants can report about how they associate the features of the rare and common diseases with their respective categories, participants were asked to estimate after each block of learning trials the conditional probabilities of the imperfect and perfect predictors given the disease, and of the disease given the imperfect and perfect predictors. We thought that estimates given early in learning, before participants had learned the objective values for the conditional probabilities, would be the most informative. If early estimates for the imperfect predictor were higher when conditional on the common disease rather than on the rare disease, this could indicate that participants were initially building up a stronger associative strength from the imperfect predictor to the common disease than to the rare disease. Also, because Kruschke (1996) proposed that the rare disease is learned by its unique symptom, it follows that the associative strength from the perfect predictor of the rare disease would be stronger than the associative strength of the perfect predictor of the common disease, suggesting that estimates of the likelihood of the rare disease conditional on its perfect predictor and the likelihood of the perfect rare predictor, conditional on the rare disease, might be perceived as greater than the corresponding quantities for the common disease. In the learning trials, both of the actual conditional probabilities for P(symptom|disease) and P(disease|symptom) were 1 for both the perfect common and perfect rare symptoms. Therefore, any early asymmetry in responses to either type of conditional probability

estimate for these perfect predictors could indicate an asymmetry in associative strength. It is important to note that using conditional probability estimates to shed light on how symptoms and diseases become associated depends on two assumptions: (1) the processes or mental representations responsible for categorization are available to consciousness and (2) assuming there is access, the conditional probability estimates are collected early enough in learning to reflect the kinds of asymmetries proposed by Kruschke (1996). In order to address concerns about early learning, data were analyzed both overall and separately for "early" and "late" learners. Also, to address some of the concerns raised in Experiments 3 and 4, Experiments 5 and 6 served to replicate the test trial data from Experiments 3 and 4 after correcting for possible methodological problems discovered with the I+PC+PR test item.

Method

Participants. Eighty-six University of Massachusetts undergraduates participated in Experiment 5 and 34 participated in Experiment 6. Experiment 5 was run during the Fall semester of 1996 and Experiment 6 was run during the Spring semester of 1997. Participants received course credit for their participation.

Stimuli and procedure. The exact stimulus structure of Experiments 2, 3, and 4 was used for the learning trials in both Experiments 5 and 6, using the same disease frequency structure as in Experiment 2 (see Table 12).

After each block of learning trials in both Experiment 5 and Experiment 6, participants made conditional probability estimates, first estimating P(disease|symptom) and then P(symptom|disease). Because there were 6 symptoms and 4 diseases there were 24 possible combinations of symptom and disease pairings for each type of conditional probability question, for a total of 48 possible estimation questions. However, to limit the number of questions, only a subset consisting of 32 questions were used, with each disease paired with the three symptoms from its own set and the imperfect predictor from the other set. In addition, following the conditional probability estimates, 40 test trials were presented after each block of Experiment 5 and 24 test trials were presented after each block of Experiment 6.

The wording of the conditional probability questions differed in Experiments 5 and 6. In Experiment 5, conditional probability questions were of the form "Given symptom x, how likely is it that patients have disease y?" or "Given disease y, how likely is it that patients have symptom x?" Participants were asked to respond on a 100-point scale, with 0 indicating no likelihood, 50 indicating equal likelihood that they have or do not have the symptom/disease, and 100 indicating certainty. Because it has been demonstrated that people have difficulty with questions that ask about probability (e.g., Myers, 1982; Hansen, McCann, & Myers, 1985), a more concrete wording of the questions expressed in terms of frequencies was used in Experiment 6. These questions were of the form "Of the next 100 patients with symptom x, how many of the patients do you think will have disease y?" or " Of the next 100 patients with disease y, how many of the patients do you think will have symptom x?".

In addition, test items similar to those used in earlier experiments were presented in each block following the learning and estimation items. The test items used in Experiments 5 were somewhat different than the test items used in Experiment 6. Both experiments used the standard (Medin & Edelson, 1988) test items I , PC, PR, PC+ PR, I+ PC+PR (see Table 30). Also, because Kruschke (1996) emphasized that ADIT is the only model that makes predictions for test items that test across disease pairs such as, I1+ PC2, I1+PR2, PC1+ PR2, and I1+PC1+ PR2, they were included in Experiment 5 for later modeling, although they were not analyzed for this thesis (see Appendix C Tables C19-C22 for results for these items). Following each block of 32 learning trials in both Experiments 5 and 6, participants were presented with test items. In both experiments, test items were presented one at a time in random order. Participants typed their responses using the keyboard, and the letters f, g, h, and j were presented on the computer screen along with the test item to remind participants of the possible disease assignments. As in prior experiments, no feedback was given during test trials.

In Experiments 3 and 4 there had been a potential problem with the counterbalancing of locations for symptoms in the I+PC+PR test item which resulted in a potential confound. In order to correct this possible problem, the order of PC and PR in the first two rows of the vertical array on the computer screen was counterbalanced while holding the position of I constant in the third row of the array, which resulted in the two orders, PC+PR+I/ PR+PC+I. Participants saw 4 PC+ PR test items, and the remainder of the test items were presented only twice for a total of 40 test items after each block in Experiment 5, or 160 total test items in Experiment 5, and 24 test trials after each block in

Experiment 6, or 96 total test items in Experiment 6. All other procedures from Experiment 3 and 4 were the same.

Table 30. Test items for Experiments 5 and 6

Test items:	
Experiment 5	Experiment 6
I	I
PC	PC
PR	PR
PC+PR	PC+PR
I+PC+PR	I+PC+PR
I+PC*	
I+PR*	
PC+PR*	
I+PC+PR*	

Note: C* is the other common disease and
R* is the other rare disease

Results-Learning and Conditional Probability Estimates

Learning Data. In Experiment 5 seven participants were dropped for not reaching the learning criterion and in Experiment 6 five participants were dropped. Participants were also classified as early and late learners. Early learners were defined as those participants who made no errors for two consecutive cycles of eight learning trials in the first two blocks of learning. Forty-six participants in Experiment 5 and 19 participants in Experiment 6 met this criterion. Late learners were defined as those participants who

reached the learning criterion during the last two blocks. Thirty-two participants in Experiment 5 and 10 participants in Experiment 6 were classified as late learners. One participant in Experiment 5 who was included in the overall analysis was dropped from the early versus late analysis because he could not clearly be classified as either an early or late learner. In Experiment 5 there was a significant difference between early and late learners for each of the first three blocks (separate variance t's block 1, $t(63)=5.72$, $p<.001$; block 2, $t(40.2)=8.22$, $p<.001$; block 3, $t(39.9)$, $p<.001$) but in Experiment 6 only block 1 and 2 were significantly different (separate variance t's block 1, $t(19.4)=3.21$; $p=.005$; block 2, $t(15.6)=5.52$, $p<.001$) and not block 3($p>.05$). Given this, it appears that learning took place at about the same rate in both Experiments 5 and 6 (see Tables 31 and 32) overall and for early learners, but that late learners in Experiment 6 learned more quickly than in Experiment 5.

Table 31. Proportion correct per block in Experiment 5

	1	2	3	4
Overall (N=79)	.71 (.02)	.87 (.02)	.94 (.01)	.97 (.01)
Early (N=46)	.79 (.02)	.96 (.01)	.97 (.01)	.98 (0)
Late (N=32)	.61 (.03)	.75 (.02)	.89 (.02)	.94 (.01)

Table 32. Proportion correct per block in Experiment 6

	1	2	3	4
Overall (N=29)	.66 (.03)	.86 (.02)	.96 (.01)	.98 (.01)
Early (N=19)	.71 (.03)	.92 (.02)	.97 (.01)	.99 (.01)
Late (N=10)	.57 (.04)	.74 (.03)	.95 (.03)	.95 (.01)

Asymmetries in Estimates for P(S|D) and P(D|S). Participants' responses to the conditional probability questions were first examined for evidence of the kinds of asymmetries suggested by Kruschke's hypotheses. Because of the concern with early learning, only results for the first two blocks are presented in detail in this section. Summary data for all four blocks are presented in Appendix C: Tables C1-C4.

Consistent with Kruschke's proposal about the expected asymmetry for the imperfect predictor, in Experiment 5 participants consistently gave significantly larger estimates for the conditional probability of the imperfect predictor given the common disease than given the rare disease across the first two blocks in the unequal frequencies condition in Experiment 5 (see Table 33). In block 1, the mean difference in estimates was 30.7 (83.4 as opposed to 52.7, $t(78)=5.81$, $p<.001$), whereas in block 2, the difference decreased to 15.4 (81.1 versus 65.7, $t(78)=2.664$, $p=.009$). This decrease of 15.3 in the difference from block 1 to block 2 was significant ($t(78)=2.35$, $p=.021$).

In Experiment 6, the pattern of results for the responses to the imperfect predictor in the unequal frequencies condition was similar to the pattern found in Experiment 5 (see Table 34). The mean difference in estimates for the imperfect predictor given the common and rare diseases in block 1 was 34.8 (72.6 compared to 37.7, t(28)=3.34, p=.002) and in block 2 the mean difference was 18.5 which was not significant (82.2 versus 63.7). As in Experiment 5, the difference in responses to the common and rare disease did decrease by 16.3 in block 2, although this interaction across blocks was not significant. There were no similar differences in the equal frequencies condition.

Table 33. Estimates for P(S|D) for the imperfect predictor-Experiment 5 N=79

	Unequal frequencies			Equal frequencies		
Block	C	R	Diff	"C"	"R"	Diff
1	83.4	52.7	30.7 ***	67.9	78.4	-10.5 *
	(3.3)	(4.3)		(3.8)	(3.5)	
2	81.1	65.7	15.4 **	72.9	74.4	-3.5
	(3.7)	(4.1)		(4.0)	(4.1)	

Note: SE is given in parentheses under the mean estimates

Table 34. Estimates for P(S| D) for the imperfect predictor-Experiment 6 N=29

	Unequal frequencies			Equal frequencies		
Block	C	R	Diff	"C"	"R"	Diff
1	72.6	37.7	34.9 **	62.8	65.4	-2.6
	(6.4)	(7.9)		(7.1)	(7.0)	
2	82.2	63.7	18.5	70.2	74.1	-3.9
	(6.0)	(8.4)		6.6)	(7.2)	

Note: SE is given in parentheses under the mean estimates

Kruschke (1996) also stated that the rare category should be encoded by its distinctive features, which implies that the symptom which is the perfect predictor for the rare disease should be more strongly associated with its disease than the perfect predictor for the common disease. Because the true conditional probabilities for P(S|D) or P(D|S) for both the common and rare perfect predictors should be 0 or 1, any asymmetry in these estimates that favored the rare disease could be interpreted as a reflection of a corresponding asymmetry in association strengths.

There was no evidence from the estimation data that there was a stronger association between the rare disease and its perfect predictor than between the common disease and its perfect predictor for either P(S|D) or P(D|S). For the estimates of P(S|D) in Experiment 5 (see Table 35) , the mean difference in block 1 in estimates was 3.6 favoring the common disease (83.6 to 80.0) and in block 2, the mean difference in estimates was 11.4 (93.9 versus 82.4, t(78)=3.32, p=.001). It should be noted that although only the difference in block 2 was significant, both differences favored the common and not the rare disease.

Table 35. Estimates for P(S| D) for the perfect predictor-Experiment 5 N=79

	Unequal frequencies				Equal frequencies		
Block	C	R	Diff		"C"	"R"	Diff
1	83.6	80.0	3.6		77.0	80.5	-3.5
	(3.3)	(3.3)			(4.0)	(3.5)	
2	93.9	82.4	11.4	***	86.2	85.7	0.5
	(1.8)	(3.5)			(2.9)	(3.3)	

Note: SE is given in parentheses next to the mean estimates

Next, the conditional probability estimates for P(D|S) for the perfect predictor in Experiment 5 were examined for any evidence of an asymmetry favoring the rare disease (see Table 36), because it was possible that an asymmetry in association strengths was present but was not apparent in the estimates for P(S|D). In contrast to Kruschke's proposal, the first two blocks showed higher estimates for the perfect predictor of the common disease (see Table 36). The average difference in response in block 1 to the common and rare diseases was 9.1 (81.0 to 71.9, t(78)=2.14, p=.04) favoring the common disease. For block 2, the average difference of 11.2 was still in the same direction (90.5 versus 79.3, t(78)= 3.24, p=.002). Because responses for both block 1 and block 2 favored the common disease, not the rare disease, there is no evidence that early associations between the perfect predictors were stronger to the rare disease than to the common disease.

Table 36. Estimates for P(D|S) for the perfect predictor-Experiment 5 N=79

	Unequal frequencies			Equal frequencies		
Block	C	R	Diff	"C"	"R"	Diff
1	81.0	71.9	9.1 *	62.0	71.1	-9.1 *
	(3.3)	(3.7)		(4.0)	(3.6)	
2	90.5	79.3	11.2 **	80.3	83.5	-3.2
	(2.2)	(3.2)		(3.2)	(3.2)	

Note: SE is given in parentheses next to the mean estimates

The conditional probability estimates of P(S|D) and for P(D|S) for the perfect predictors of the rare and common diseases were also examined in Experiment 6. These results were similar to those in Experiment 5 (see Tables 37 and 38). There was no

evidence to show that participants showed a stronger association for the perfect predictor of the rare disease than for the perfect predictor of the common disease. For the estimates of P(S|D), the mean difference in block 1 was 8.58 in favor of the common disease (66.48 to 57.90) and the mean difference in block 2 was 2.17 (91.38 versus 89.21). For the estimates of P(D|S) in block 1, the mean difference in estimates was 4.31 favoring the common disease (68.17 to 63.86) and in block 2 the mean difference in estimates was -1.76 (87.79 versus 89.55). Although none of the conditional probability estimates in Experiment 6 significantly favored the common disease, three of four differences were positive, and both block 1 estimates were larger for the common disease than for the rare disease. Taken together, these results provided no evidence that there was a stronger association for the perfect rare symptom to its category than for the perfect common symptom to its category. In fact, the results from Experiment 5 suggest that the early association might have been stronger for the perfect common symptom to its disease than that of the perfect rare symptom.

Table 37. Estimates for P(S| D) for the perfect predictor-Experiment 6 N=29

	Unequal frequencies			Equal frequencies		
Block	C	R	Diff	"C"	"R"	Diff
1	66.48	57.90	8.58	72.59	64.31	8.28
	(6.8)	(6.7)		(6.8)	(7.7)	
2	91.38	89.21	2.17	89.55	89.21	0.34
	(4.4)	(4.9)		(4.6)	(5.0)	

Note: SE is given in parentheses next to the mean estimates

Table 38. Estimates for P(D|S) for the perfect predictor-Experiment 6 N=29

	Unequal frequencies			Equal frequencies		
Block	C	R	Diff	"C"	"R"	Diff
1	68.17	63.86	4.31	72.83	66.24	6.59
	(7.3)	(6.6)		(6.2)	(6.8)	
2	87.79	89.55	-1.76	82.28	77.21	5.07
	(5.4)	(4.4)		(5.3)	(6.6)	

Note: SE is given in parentheses next to the mean estimates

It should be noted that overall conditional probability estimates were similar in Experiments 5 and 6. This is evidence that that the different wordings used for the conditional probability questions in Experiments 5 and 6 did not have a great impact on response patterns, even though previous work (e.g., Myers, 1982; Hansen, McCann, & Myers, 1985) had suggested that they might.

Summary-Overall Estimation Data. For estimates of P(S|D) for the imperfect predictor in the unequal early condition, where the true conditional probability was 1, participants in both experiments responded with larger estimates for the common disease than for the rare disease. This is consistent with the association strengths hypothesized by Kruschke(1996). There was not, however any direct evidence for Kruschke's proposal that the rare disease should be encoded by its distinctive feature. Evidence for this would have been found in any indication that the association between the rare disease and its perfect predictor was stronger than the corresponding association for the common disease. None of the data from Experiments 5 and 6 showed higher estimates for the rare disease than for the common disease as would be expected if Kruschke's hypothesis was correct.

Asymmetries in Estimates for P(S|D) and P(D|S) for Early and Late Learners. Conditional probability estimates were also examined separately for early and late learners in both Experiments 5 and 6. There were a number of reasons we were interested in looking at early and late learners. First, looking at late learners in the first two blocks might give a better picture of what occurs early in learning. That is, any asymmetry that might exist early on in learning might be more readily apparent in the data of late learners than in the data of early learners. We also might find evidence of an interaction of early and late learning with the imperfect and perfect predictors for the rare and common diseases. This is relevant because specific asymmetries in associations between features and categories have been proposed as underlying the inverse base-rate effect. Therefore, if these asymmetries differ for early and late learners at any point in learning, this might influence whether or not the inverse base-rate effect develops at different points in learning for either disease pair. That is, specific asymmetries in associations might co-occur with the inverse base-rate effect.

First, the response patterns for P(S|D) were examined for the imperfect predictor for early and late learners in Experiments 5 and 6 (see Table 39). In Experiment 5 there were no differences in response patterns for early and late learners across the first two blocks in responses to the imperfect predictor in the unequal early and unequal late conditions, and both early and late learners responded in the same pattern as was found in the overall data. Both early and late learners gave higher estimates for the conditional probability of the imperfect predictor given the common disease than for the imperfect predictor given the rare disease (block 1- $t(45)=5.31$, $p<.001$; $t(31)=2.71$, $p=.01$- early and late respectively; block 2 –$t(45)=2.43$, $p=.02$- early). In Experiment 6, estimates of

P(S|D) for the imperfect predictor for both early and late learners were similar to those found in Experiment 5. In Experiment 6 the differences for early and late learners in block 1 and 2 in the unequal early condition were 34.63 and 35.30, (t(18)=2.64, p=.02; t(9)=1.94, p=.08) and in block 2- 28.31 (early-t(18)=2.19, p=.04) and 0 (late). There was no such asymmetry in the unequal late condition and, in addition, no differences between early and late learners were found for the last two blocks in either experiment. Both early and late learners' data in Experiments 5 and 6 were consistent with Kruschke's hypothesis that the asymmetry in association for the imperfect predictor to the common and rare diseases should favor the common disease.

Tables 39. P(S(imperfect)|D) for the unequal early condition –early versus late learners – Experiments 5 and 6

	Experiment 5						Experiment 6					
	Early learners			Late learners			Early learners			Late learners		
Block	C	R	Diff	C	R	Diff	C	R	Diff	C	R	Diff
1	86.96	54.00	35.96***	77.84	52.47	25.37*	69.21	34.58	34.63*	79.00	43.7	35.3
	(3.74)	(6.04)		(6.19)	(6.20)		(8.37)	(9.64)		(10.05)	(14.13)	
2	83.52	68.65	14.87*	76.97	63.41	13.56	86.68	58.37	28.31*	73.70	73.70	0
	(4.11)	(5.22)		(6.83)	(6.55)		(6.19)	(10.84)		(12.84)	(12.89)	

Note: SE is given in parentheses under the mean estimates

For estimates of P(S|D) for perfect predictors in the unequal early condition in Experiment 5 there were few differences found in the response patterns for early and late learners (see Table 40). In block 1 and 2 for early learners and block 1 for late learners there were no significant differences in estimates of the perfect predictor favoring either

the common or rare diseases. In the one instance where a bias in estimates did exist it was a bias for the common diseases (block 2 late learners: mean difference of 21.66- $t(9)=3.84$, $p=.001$). Again, Experiment 6 data (see Table 40) for the P(S|D) for the perfect predictors in the unequal early condition was remarkably similar to that of Experiment 5 for early and late learners and to the overall data for Experiments 5 and 6. For blocks 1 and 2 for both early and late learners there were no significant differences in estimates that favored either the common or the rare disease. This pattern of responses found for both types of learners reflected the pattern found in the overall data (see estimation data) and indicated that, for both early and late learners, there were no asymmetries favoring the rare perfect predictor of the type predicted by Kruschke (1996).

Tables 40. P(S(perfect)|D) for the unequal early condition-early versus late learners- Experiments 5 and 6

	Experiment 5						Experiment 6					
	Early learners			Late learners			Early learners			Late learners		
Block	C	R	Diff	C	R	Diff	C	R	Diff	C	R	Diff
1	86.70	88.91	-2.21	78.53	66.97	11.56	64.26	64.21	0.05	70.70	45.90	24.80
	(3.74)	(3.74)		(6.12)	(5.49)		(8.91)	(7.42)		(10.61)	(12.98)	
2	91.89	87.24	4.65	96.50	74.84	21.66***	92.53	85.58	6.95	89.20	96.10	-6.90
	(2.74)	(3.97)		(1.88)	(6.27)		(4.96)	(7.42)		(8.86)	(1.95)	

Note: SE is given in parentheses under the mean estimates

Next examined were responses to the P(D|S) estimates for the perfect predictors for early and late learners in Experiments 5 and 6 (see Table 41). First, in Experiment 5 there was a type of learner by response interaction in blocks 1 and 2 in the unequal early condition ($F(1,76)=4.86$, $p=.03$; $F(1,76)=17.23$, $p<.001$ for blocks 1 and 2 respectively).

Although the objective conditional probability for the perfect predictor of either the common or rare disease is 1, late learners showed a consistent bias, responding with substantially lower responses for the rare disease probe than for the common disease probe (78.53 to 58.03 and 95.72 to 68.44). Early learners, however responded with similar estimates for both common and rare test items for the perfect predictors (82.37 to 80.85 and 86.63 to 86.48). Neither the pattern for early learners in Experiment 5 nor the pattern for late learners in Experiments 5 and 6 are consistent with the Kruschke (1996) hypothesis that the perfect predictor of the rare disease will be more associated to its disease than the perfect predictor of the common disease. In fact, the pattern for the late learners is the opposite of that predicted by Kruschke (1996) and is consistent with the pattern found in the overall estimation data. It should be noted that in the case of late learners in Experiment 6 the response pattern is in the direction predicted by Kruschke (1996), providing some evidence that for late learners the perfect predictor for the rare disease is more closely associated with its disease than the perfect predictor of the common disease is associated with its disease.

Finally, we examined the responses of early and late learners to P(D|S) for the imperfect predictors in Experiment 5 and Experiment 6. Because an asymmetry is predicted for the true conditional probabilities, these items were not examined in the collapsed data. They were examined here, however, for any effect of early and late learning, or any interaction of early and late learning with response patterns to examine whether early and late learners differ in developing the inverse base-rate effect because of differences in their associations of the imperfect predictor to the rare and common

diseases. In Experiments 5 and 6 (see Table 42), we found no significant differences for early and late learners in their responses to the imperfect predictors.

Tables 41. P(D|S(perfect)) for the unequal early condition-early versus late learners- Experiments 5 and 6

	Experiment 5						Experiment 6					
	Early learners			Late learners			Early learners			Late learners		
Block	C	R	Diff	C	R	Diff	C	R	Diff	C	R	Diff
1	82.37	80.85	1.52	78.53	58.03	20.50**	72.68	56.58	16.10	59.60	77.70	-18.10
	(3.83)	(4.47)		(6.16)	(5.59)		(8.31)	(8.02)		(14.25)	(10.92)	
2	86.63	86.48	0.15	95.72	68.44	27.28***	84.16	91.05	-6.89	94.70	86.70	8.00
	(3.44)	(3.44)		(1.78)	(5.79)		(7.99)	(5.72)		(2.74)	(6.74)	

Note: SE is given in parentheses under the mean estimates

Tables 42. P(D|S(imperfect)) for the unequal early condition-early versus late learners- Experiments 5 and 6

	Experiment 5							Experiment 6					
	Early learners				Late learners			Early learners			Late learners		
Block	C	R	Diff		C	R	Diff	C	R	Diff	C	R	Diff
1	70.46	45.09	25.37	***	66.62	49.38	17.24	66.58	34.68	31.90 *	52.40	27.80	24.60
	(5.03)	(5.20)			(6.65)	(6.47)		(8.76)	(8.83)		(13.41)	(9.90)	
2	69.89	56.59	13.30	*	65.91	55.62	10.29	70.63	50.42	20.21	72.20	65.20	7.00
	(4.58)	(5.06)			(7.00)	(6.28)		(6.58)	(9.24)		(12.95)	(13.45)	

Note: SE is given in parentheses under the mean estimates

Summary-Estimation Data for Early and Late Learners. Based on the estimates of P(S|D) and P(D|S) for the imperfect predictor in the unequal early condition, any differences in the development of the inverse base-rate effect for early and late learners cannot be attributed to differences found in reported

associations from the imperfect predictor to the diseases or from the diseases to the imperfect predictor. If this were the case we would have expected to find different response patterns for early and late learners in either experiment, but no such differences were found.

We also found little evidence in estimates for the P(S|D) or P(D|S) for the perfect predictors for Kruschke's (1996) hypothesis that the rare disease is encoded by its distinctive feature for either early or late learners. Although late learners in Experiment 6 did show the asymmetry predicted by Kruschke in block 1 it was not significant and was the only evidence in all of the estimation data for an asymmetry of the type predicted by Kruschke.

Accuracy in Estimates for P(S|D) and P(D|S) for Early and Late Learners. Finally, the responses for P(S|D) and P(D|S) for "distractor" items for both Experiments 5 and 6 were analyzed. Distractor items were those items for which the objective values of P(S|D) and P(D|S) were 0 and were included as a test of accuracy. For P(S|D) in Experiment 5 (see Appendix C Tables C5-C8), there was an effect of early and late learning on the estimates for blocks 1 and 2 ($t(76)=5.21$, $p<.001$; $t(76)=3.19$, $p=.002$), where early learners made significantly lower estimates than did the late learners. There was a similar effect in Experiment 6, also in blocks 1 and 2 ($t(27)=4.20$, $p<.001$; $t(27)=2.69$, $p=.01$). By blocks 3 and 4, for both experiments, there were no significant differences in the responses to distractor items for early and late learners.

The estimates for P(D|S) were also examined for any differences in responses to the distractor items for early and late learners. Again, in Experiment 5, there were significantly higher estimates made by late learners to the distractor items than for the

estimates made by early learners to the same items. This effect of early and late learning was highly significant in the first two blocks ($t(76)=3.90$, $p<.001$; $t(76)=4.41$, $p<.001$). The results for Experiment 6 showed the same effect for blocks 1 and 2 ($t(27)=4.37$, $p<.001$; $t(27)=3.66$, $p=.001$). By blocks 3 and 4 in Experiment 5 there was only a marginally significant difference for early and late learners, and there was no difference in Experiment 6.

These results indicate that early learners are initially more accurate (i.e. are giving estimates closer to the objective probability of 0) than late learners in both Experiment 5 and Experiment 6, but that by the end of learning, accuracy is the same for both early and late learners.

Results-Test Items

Even though the primary motivation for Experiments 5 and 6 was to collect conditional probability estimates, test items were also presented after each block. This was done to replicate the work of Experiment 3 and to correct the potential problem with the I+PC+PR test item.

The Inverse Base-rate Effect. Of greatest interest were responses to the PC+ PR test item. The results for Experiments 5 and 6 were similar to the pattern of results found in Experiments 3 and 4. As can be seen in Table 43, in Experiment 5 the inverse base-rate effect was present in the overall data in the unequal early condition by the end of block 2, and remained present for blocks 3 and 4 ($t2(79)= -2.25$, $p=.03$; $t3(79)= -2.96$, $p=.004$; $t4(79)= -3.11$, $p=.003$). For Experiment 6 (see Table 44), participants showed an

overall trend towards an inverse base-rate in the unequal early condition across all 4 blocks, but the effect was only significant in block 3 (t3(28)= -3.02, p=.005), perhaps due to the smaller number of participants. There was no inverse base-rate effect in the unequal late condition for either Experiment 5 or Experiment 6.

Table 43. Response proportions to test item PC+PR-Experiment 5

		Unequal early				Unequal late			
	Block	C	R	Diff		C	R	Diff	
Overall	1	.38	.51	-.13		.46	.39	.07	
	2	.37	.58	-.21	*	.46	.47	-.01	
	3	.36	.61	-.25	**	.43	.50	-.07	
	4	.35	.60	-.25	**	.46	.51	-.05	
Early (N=46)	1	.30	.62	-.32	**	.48	.40	.08	
	2	.32	.67	-.35	**	.49	.47	.02	
	3	.29	.68	-.39	***	.49	.46	.03	
	4	.30	.68	-.38	***	.53	.46	.06	
Late (N=32)	1	.48	.36	.12		.44	.38	.06	
	2	.45	.45	0		.41	.48	-.07	
	3	.45	.50	-.05		.34	.58	-.24	****
	4	.42	.48	-.06		.34	.59	-.25	****

Note: **** indicates significance when results are collapsed across blocks 3 and 4

The Inverse Base-rate Effect in Early and Late Learners. In analyzing early and late learners separately, there were some interesting results, again similar to the results found in Experiments 3 and 4. In the unequal early condition in Experiment 5, early learners showed a consistent inverse base-rate effect across all four blocks (t1(45)= -2.92,p=.006; t2(45)= -3.18, p=.003; t3(45)= -3.60, p=.001; t4(45)= -3.49, p=.001), even though the diseases were presented with equal frequencies during the last two learning blocks. However, in the unequal late condition, there was no difference for early learners

in responses to the PC+PR test item. Both of these results indicate that early learners only exhibited a sensitivity to the early underlying frequencies.

Table 44. Response proportions to test item PC+PR-Experiment 6

		Unequal early				Unequal late		
	Block	C	R	Diff		C	R	Diff
Overall	1	.33	.45	-.12		.43	.48	-.05
	2	.38	.59	-.21		.48	.50	-.02
	3	.31	.69	-.38	***	.50	.50	0
	4	.36	.64	-.28		.50	.48	.02
Early (N=19)	1	.37	.47	-.10		.45	.50	-.05
	2	.37	.63	-.26		.50	.50	0
	3	.31	.68	-.37	*	.50	.50	0
	4	.34	.66	-.32		.50	.50	0
Late (N=10)	1	.25	.40	-.15		.40	.45	-.05
	2	.40	.50	-.10		.45	.50	-.05
	3	.30	.70	-.40		.50	.50	0
	4	.40	.60	-.20		.50	.50	0

Also, as in Experiment 3, late learners in Experiment 5 did not show an inverse base-rate effect in the unequal early condition at any point in learning. However, they did start to show a slight inverse base-rate effect in the unequal late condition when the disease frequencies become unequal (collapsing across blocks 3 and 4: $t(63)= -2.56$, $p=.01$). This suggests that the late learners exhibited a sensitivity to the underlying frequencies in blocks 3 and 4. This, along with the data for early learners is in accord with our hypothesis that association strengths are influenced most strongly by the frequency patterns present during learning and continue to be influenced by frequencies if learning is still occurring. In Experiment 6 there is an inverse base-rate trend in the unequal early condition for both early and late learners throughout all four blocks,

although it is only significant in block 3 for early learners($t(18) = -2.348$, $p=.03$). Early and late learners do not show any differences in response patterns in either the unequal early or unequal late condition. The differences in results from Experiments 5 and 6 for the PC+PR test item could be attributed to two factors. First there may not have been a large enough group run in Experiment 6 to separate into early and late learners. Because there were only 19 early and 10 late learners, there was little power to find such a difference. Secondly, when we examined the learning data from early and late learners in both experiments (see learning data) the difference between the two groups appeared to continue throughout block 3 in Experiment 5 while it dissipated by block 3 in Experiment 6, indicating that there might have been a greater difference between early and late learners in Experiment 5 than there was in Experiment 6.

The I+PC+PR Triplet. Next we examined the results for the I+PC+PR item. In Experiment 5 (see Table 45) as in Experiment 3, in the unequal early condition there was a trend towards a positive base-rate bias for the first two blocks, but this was only marginally significant in block 2 ($t(77)=2.079$, $p=.04$).This overall trend disappeared by the third block. This was similar to the pattern found in the collapsed data in Experiment 6, but the positive base-rate bias was only significant in block 1 ($t(28)=2.781$, $p=.003$). In the unequal late condition for blocks 1 and 2 there was no reason to expect a bias in responding favoring either disease and none of the results for Experiment 5 or Experiment 6 were significant overall.

The I+PC+PR Triplet for Early and Late Learners. Highly interesting in Experiment 5 was a type of learner by response by frequency interaction collapsed across

the final two blocks ($F(1,154)=9.32$, $p=.003$). That is, for blocks 3 and 4, early learners responded with a negative bias in the unequal early condition, as they do in Experiment 3, and with a positive bias in the unequal late condition (.43 to .55 and .39 to .59; .55 to.42 and .53 to.44; $t(91)=-2.37$, $p=.02$), whereas, late learners showed the opposite pattern and responded with a negative bias in the unequal late condition and a positive bias in the unequal early condition(.36 to .56 and .43 to.52; .59 to.41 and .51 to .41; $t(63)=2.00$, $p=.05$). Because the presentation order was couterbalanced properly in Experiments 5 and 6 due to the potential confound in Experiments 3 and 4, it seems as though there truly is an inverse triplet base-rate effect. This, along with the inverse base-rate data for early and late learners, provides additional evidence that response patterns are influenced by the frequency patterns present throughout learning and can change over the course of learning if learning is still taking place, as it would be in the case of late learners.

As in the overall data, in the unequal early condition, there was also a bias toward the more common disease for late learners in blocks 1 and 2 in Experiment 5 ($t(31)=2.63$, $p=.01$, $t(31)=2.82$, $p=.008$, blocks1 and 2 respectively) which was also found in the data of early learners in blocks 1 and 2 in Experiment 6 ($t(18)=2.44$, $p=.02$, $t(18)=2.35$, $p=.03$, blocks 1 and 2 respectively). It must be noted that there was one minor anomaly in Experiment 5, where there was a marginally significant inverse pattern of responding found for late learners in the unequal late condition prior to the frequency shift in block 2 ($t(31)= -2.31$, $p=.03$). However, it is possible that this was due to the 11% of responses that late learners gave to inappropriate targets in this block.

Table 45. Response proportions to test item I+PC +PR for Experiment 5

		Unequal early				Unequal late			
	Block	C	R	Diff		C	R	Diff	
Overall	1	.50	.38	.12		.38	.50	-.12	
	2	.57	.40	.17	*	.39	.54	-.15	
	3	.49	.49	0		.47	.48	-.01	
	4	.44	.52	-.08		.49	.47	.02	
Early	1	.46	.46	0		.43	.50	-.07	
	2	.52	.46	.06		.45	.52	-.07	
	3	.43	.55	-.12		.55	.42	.13	
	4	.39	.59	-.20		.53	.44	.09	
Late	1	.59	.26	.33	**	.29	.50	-.21	
	2	.64	.26	.33	**	.30	.59	-.29	*
	3	.59	.41	.18		.36	.56	-.20	
	4	.51	.41	.10		.43	.52	-.09	

Table 46. Response proportions to test item I+PC+PR for Experiment 6

		Unequal early				Unequal late		
	Block	C	R	Diff		C	R	Diff
Overall	1	.66	.24	.42	*	.52	.38	.14
	2	.63	.37	.26		.47	.49	-.02
	3	.49	.51	-.02		.51	.47	.04
	4	.42	.58	-.16		.47	.52	-.05
Early	1	.70	.26	.44*	*	.59	.37	.22
	2	.68	.32	.36*	*	.47	.53	-.06
	3	.53	.47	.06		.61	.39	.32
	4	.42	.58	-.16		.54	.45	.09
Late	1	.58	.20	.38		.38	.40	-.02
	2	.52	.48	.04		.48	.42	.06
	3	.42	.58	-.16		.32	.62	-.30
	4	.42	.58	-.16		.32	.65	-.33

The Imperfect Predictor. For the test item for the imperfect predictor (I) for Experiment 5 the response pattern was similar to the pattern found in previous experiments (see Experiments 3 and 4). Participants showed a positive base-rate bias for all four blocks for the pair of diseases in the unequal early condition (see Table 47) which showed correct use of base-rate information. This pattern was present for both early and late learners, although for late learners the effect did not become significant until block 3. A similar pattern was observed in the data from Experiment 6 (see Table 52). In Experiment 5, in the unequal late condition there was an anomalous finding. Although there is no reason in blocks 1 and 2 for a bias to exist for the imperfect predictor there was an significant effect favoring the disease that later became the rare disease both overall and for early learners (overall: block 1-$t(77) = -2.94$, $p=.004$; block 2-$t(77) = -2.25$,

p=.03 early: block 1-t(45)= -2.58, p=.01; block 2-t(45)= -1.97, p=.05). There were no biases in the unequal late condition favoring either the common or rare disease for late learners in Experiment 5 or in Experiment 6 both overall and separately for early and late learners.

Table 47. Response proportions to the imperfect predictor-Experiment 5

		Unequal early				Unequal late			
	Block	C	R	Diff		C	R	Diff	
Overall	1	.61	.21	.40	***	.29	.54	-.25	**
	2	.66	.21	.45	***	.35	.56	-.21	*
	3	.70	.25	.45	***	.41	.54	-.13	
	4	.67	.26	.41	***	.40	.53	-.13	
Early	1	.73	.14	.59	***	.32	.62	-.30	*
	2	.76	.16	.60	***	.36	.61	-.25	*
	3	.74	.23	.51	***	.42	.55	-.13	
	4	.70	.24	.46	***	.42	.55	-.13	
Late	1	.44	.31	.13		.25	.41	-.16	
	2	.52	.28	.24		.34	.50	-.16	
	3	.66	.28	.38	*	.39	.53	-.14	
	4	.64	.30	.34	*	.36	.48	-.12	

Table 48. Response proportions to the imperfect predictor-Experiment 6

		Unequal early				Unequal late		
	Block	C	R	Diff		C	R	Diff
Overall	1	.62	.28	.34	*	.41	.33	.07
	2	.74	.26	.38	**	.33	.53	-.20
	3	.71	.28	.43	**	.47	.52	-.05
	4	.67	.33	.34	*	.40	.60	-.20
Early	1	.68	.26	.42	*	.50	.34	.16
	2	.72	.29	.43	*	.37	.58	-.21
	3	.66	.32	.34		.37	.61	-.24
	4	.74	.26	.48	**	.37	.63	-.26
Late	1	.50	.30	.20		.25	.30	-.05
	2	.80	.20	.60	*	.25	.45	-.20
	3	.80	.20	.60	*	.65	.35	.30
	4	.55	.45	.10		.45	.55	-.10

The Perfect Common and Perfect Rare Predictors. As in Experiments 3 and 4, participants in Experiments 5 and 6 overall learned the correct assignment of the perfect predictor of the common disease (PC) in the unequal early condition by the end of block 1 (Experiment 5 block 1-$t(77)=9.47$, $p<.001$; Experiment 6 block 1-$t(28)=3.89$, $p<.001$; see Tables 49 and 50). This pattern was present for both early learners (Experiment 5 block1 $t(45)=11.66$, $p<.001$; Experiment 6 block 1 $t(18)=3.76$, $p=.001$) and late learners (Experiment 5 block1-$t(31)=3.53$, $p=.001$; Experiment 6 block 2-$t(9)=4.58$, $p=.001$). Participants in Experiments 5 and 6 learned the correct responses for the perfect rare predictor (PR), by the end of block 2 (Experiment 5 block 1-$t(77)=-6.84$, $p<.001$; Experiment 6 block 2-$t(28)=-7.08$, $p<.001$- see Tables 51 and 52) This pattern of responses was present for both early learners (Experiment 5 block1-$t(45)=-13.86$, $p<.001$; Experiment 6 block1-$t(18)=-3.50$, $p=.003$) and late learners (Experiment 5 block 2-$t(31)=-4.19$, $p<.001$; Experiment 6 block2- $t(9)=-6.00$, $p<.001$).

Table 49. Response proportions to the perfect common predictor-Experiment 5

		Unequal early			Unequal late		
	Block	C	R		C	R	
Overall	1	.78	.13	***	.58	.16	***
	2	.90	.05	***	.78	.08	***
	3	.91	.06	***	.87	.03	***
	4	.95	.03	***	.94	.01	***
Early	1	.86	.08	***	.72	.06	***
	2	.94	.03	***	.90	.03	***
	3	.92	.05	***	.95	0	***
	4	.97	.03	***	.99	0	***
Late	1	.67	.20	***	.39	.30	***
	2	.86	.08	***	.59	.16	***
	3	.89	.06	***	.77	.18	***
	4	.92	.03	***	.88	.02	***

Table 50. Response proportions to the perfect common predictor-Experiment 6

		Unequal early			Unequal late		
	Block	C	R		C	R	
Overall	1	.66	.19	**	.53	.28	
	2	.81	.17	***	.91	.03	***
	3	.83	.12	***	.98	0	***
	4	.90	.03	***	.97	.03	***
Early	1	.71	.13	***	.61	.26	
	2	.79	.18	**	.92	.03	***
	3	.87	.08	***	1.00	0	***
	4	.87	.05	***	.97	.03	***
Late	1	.55	.30		.40	.30	
	2	.85	.15	***	.90	.05	***
	3	.75	.20	*	.95	0	***
	4	.95	0	***	.95	.05	***

Table 51. Response proportions to the perfect rare predictor-Experiment 5

		Unequal early			Unequal late		
	Block	C	R		C	R	
Overall	1	.15	.64	***	.21	.63	***
	2	.08	.83	***	.08	.84	***
	3	.03	.89	***	.04	.92	***
	4	.01	.92	***	.03	.91	***
Early	1	.04	.84	***	.14	.79	***
	2	.01	.96	***	.04	.96	***
	3	0	.94	***	0	.99	***
	4	0	.97	***	.01	.97	***
Late	1	.30	.36		.31	.39	
	2	.17	.66	***	.12	.67	***
	3	.06	.83	***	.09	.83	***
	4	.03	.84	***	.06	.83	***

Table 52. Response proportions to the perfect rare predictor-Experiment 6

		Unequal early			Unequal late		
	Block	C	R		C	R	
Overall	1	.17	.43	*	.29	.55	
	2	.09	.84	***	.14	.86	***
	3	0	.97	***	.05	.95	***
	4	.02	.95	***	.03	.97	***
Early	1	.13	.63	**	.26	.63	*
	2	.13	.87	***	.13	.87	***
	3	0	1.00	***	.05	.95	***
	4	.03	.95	***	0	1.00	***
Late	1	.25	.05		.35	.40	
	2	0	.80	***	.15	.85	*
	3	0	.90	***	.05	.95	***
	4	0	.95	***	.10	.90	***

Summary-Test Item Data

Experiment 5, like Experiments 3 and 4, again showed us that the inverse base-rate effect developed in the unequal early condition quite early in learning in the overall data and for early learners, and this pattern remained present throughout learning. However, late learners in Experiment 5 did not exhibit an inverse base-rate effect in the unequal early condition, but showed an inverse base-rate effect in the unequal late condition. This again provided evidence that frequencies continued to exert an influence on stored representations as long as learning was ongoing, but stopped having an influence once learning had been completed.

For the I+PC+PR triplet, the most intriguing result was the comparison of early and late learners in Experiment 5, where there was additional evidence for the ongoing

influence of frequencies on stored representations, contingent upon the frequencies during which learning took place. Whereas early learners showed what we will call an inverse triplet base-rate effect in the unequal early condition and a positive base rate effect in the unequal late condition, the late learners showed the inverse triplet base-rate effect in the unequal late condition and a positive base rate-effect in the unequal early condition. This reconfirms the findings from Experiments 3. It is important to note that previous research has not found such an inverse triplet base-rate effect, but it was possible that the effect was present but was not discovered because results were averaged over early and late learners.

Finally, there was evidence that the imperfect predictor was more associated with the common disease, for both overall and early and late learners in both Experiments 5 and 6.

CHAPTER 6

GENERAL DISCUSSION

Conclusion

Most research on the inverse base-rate effect in the past has focused on the category representations that are formed once learning is completed. In these experiments our goals were fourfold: (1) to demonstrate a robust inverse base-rate effect (2) to examine the time course of learning (3) to see how frequency patterns affected the category representation formed throughout the course of learning and (4) to specifically examine Kruschke's (1996) proposal that the inverse base-rate effect is caused by the rare disease being learned in contrast to the common disease.

First, we were able to replicate the inverse base-rate effect in Experiment 1 using a standard experimental design. Hopefully, this will lay to rest some of the questions that have arisen as to whether or not the inverse base-rate effect truly exists. We also were able to more clearly interpret findings about the influence of frequency patterns on learning by using a simplified version of the Medin & Bettger (1991) study, first modifying the design in Experiment 2, and then by probing participants throughout the course of learning in Experiments 3, 5, and 6, and by dividing learners into early and late learners in Experiments 3 through 6. These studies showed that it is possible to make a valid distinction between early and late learners, while still requiring all participant to reach a certain level of performance by the end of the learning trials. All of the learning

data from our experiments indicated that by the end of learning, early learners were not distinguishable from late learners and yet they showed distinctly different patterns of responding. This level of analysis is interesting in that it is not simply looking at the aggregate data, where you lose the applicability to the individual participant by looking at averages, but it also does not suffer from the problems that arise when looking at individual differences, where it is difficult to generalize from any one individual to another. This type of analysis provides a distinction somewhere in-between these two levels by looking at two subpopulations and can potentially be a useful tool for researchers to employ when analyzing data.

In terms of specific questions about the development of the inverse base-rate effect, we were able to clearly show evidence for the early development of the inverse base rate-effect (by the end of the second block of learning) in Experiments 3 and 5, and we also demonstrated that learning must occur fairly early on for the inverse base-rate effect to develop, as it did not develop for late learners in either Experiment 3 or 5. However, we did find that late learners showed an inverse base-rate effect in the unequal late condition in Experiment 5, and we took this as new evidence for the influence of later frequencies. We feel this is an indication that the category representation formed is malleable as long as learning is ongoing, and can continue to be influenced by frequency patterns until learning has been completed. This has implications for Kruschke's (1996) ADIT model. Kruschke(1996) stated that "people consistently favor the more frequent category" and that any "apparent inconsistencies in base-rate utilization observed in the inverse base-rate effect are merely apparent; people are consistently applying base-rate knowledge to asymmetric category representations." Based on our results, this would

need to be qualified by the statement that people favor the category that is more frequent while learning is ongoing up to the point where learning is completed. Therefore, the asymmetry in category representation that Kruschke(1996) posits to cause the inverse base-rate effect will differ depending on when learning is completed. Also, as stated earlier, ADIT learns more like late learners than early learners. Because late learners in Experiments 3 and 5 never show the inverse base-rate effect in the unequal early condition it seems unlikely, given the current evidence, that the mechanisms driving ADIT, although they are psychologically plausible, are the mechanisms that lead to the development of the inverse base-rate effect.

Given the differences in the development of the inverse base-rate effect for early and late learners we hoped to find evidence of different asymmetries in associations of features to categories in the conditional probability estimates. However, we were unable to find any such differences. We did find additional evidence for the influence of late frequencies on early learners in responses to the triplet test item in Experiment 3, and on both early and late learners for the same test item in Experiment 5. The inverse triplet base-rate effect found in the unequal early condition for early learners and in the unequal late condition for late learners in the last two blocks in Experiment 5 seemed to parallel the development of the inverse base-rate effect for early learners in the unequal early condition and late learners in the unequal late condition. This indicates that even in the situation where all the features of each disease category are present, and where previous researchers have indicated that decisions are based on base rate information, the story is much more complicated than it first appears to be. Any model now will have to account

for both the pattern of responses for early learners and the pattern of responses for late learners for both the inverse base-rate effect and for the inverse triplet base-rate effect.

Finally, we also examined Kruschke's proposal that participants first learn the common category and then learn the rare category in terms of its distinctive features. We found evidence in the responses to the imperfect predictor in Experiments 3, 4, 5, and 6 and in conditional probability results in Experiments 5 and 6 for the first part of Kruschke's principle that the common category is considered typical and that the imperfect predictor will therefore be more associated with the common category than with the rare category. However, we could find no evidence in the conditional probability estimates in Experiments 5 and 6 for the proposal that the rare category is encoded primarily by its distinctive features. Once again, it should be restated that using conditional probabilities as estimates of the associations between categories and features of categories depends on the assumptions that (1) the associations that were formed were at an accessible level of consciousness and (2) that the estimates were collected early enough in learning to detect any asymmetries in associations that may be present.

Some Additional Questions For Future Research

There are additional questions that might be addressed in future research. First, can an analog of the inverse base-rate effect occur in a situation with equal frequencies by causing one disease to be learned prior to another (Kruschke,1996) ? One possible way to cause one category to be learned prior to another while still maintaining equal frequencies is to systematically manipulate the order in which cases are presented within a block of

trials. Another possible way of influencing category learning is to have participants utilize what they have learned, for example by having them prescribe a drug based on their diagnosis. There is evidence with this type of manipulation, an analog of the inverse base-rate effect can be obtained (Ross, psychonomics presentation, 1996). Until this is tested, there will not be any solid evidence for Kruschke's (1996) proposal that inverse responding is not uniquely caused by unequal frequencies, but may be caused by other factors, such as learning one item before another. Also, do the models (Kruschke,1996; Myers, Lohmeier, & Well,1994) perform in the same way as people? It will be important to test both the revised CLEM model (Myers, Lohmeier, & Well,1994) and ADIT model (Kruschke,1996) to see if they can account for the full set of the above results.

Finally, it appears as though there is a different allocation of attentional resources to symptoms in diseases with unequal frequencies corresponding to different attention to features of categories and that human differences in attentional allocation should correspond with the attentional parameter in ADIT (Kruschke,1996). Is there a way to test human attention directly, for example by monitoring how much time participants spend looking at each symptom? If so, do shifts in attention in humans lead to the same responses as shift of attention in ADIT (Kruschke,1996)?

In summary, the experiments presented in this thesis did help to determine some of the causal agents involved in producing the inverse base-rate effect. Kruschke's model appeared on the surface to be psychologically plausible, and the above studies provided a starting point to further examine his claims and provided preliminary evidence that they may not be the mechanisms that are operating. However, there is still much work that needs to be done to understand the exact psychological mechanisms underlying the effect.

APPENDIX A

TABLES FOR EXPERIMENT 2

Table 53. Complete response proportions for Experiment 2

	Unequal early		Unequal late	
Test Items	C1	R1	C2	R2
I1	.81	.15	.01	.03
PC1	.96	0	0	.04
PR1	0	.96	.03	.01
PC1+PR1	.40	.60	0	0
I1+PC1+PR1	.61	.38	.01	0
I1+PC2	.18	.11	.69	.01
I1+PR2	.21	.04	.01	.65
PC1+PR2	.33	.01	0	.65
I1+PC1+PR2	.83	.01	.01	.14
I2	0	.10	.50	.40
PC2	0	.01	.93	.06
PR2	.01	.04	.01	.93
PC2+PR2	.04	0	.44	.51
I2+PC2+PR2	.01	.01	.56	.42
I2+PC1	.69	.01	.14	.15
I2+PR1	.01	.72	.15	.11
PC2+PR1	0	.57	.43	0
I2+PC2+PR1	0	.36	.57	.07

APPENDIX B

TABLES FOR EXPERIMENTS 3 AND 4

Table 54. Complete data for test item I-Experiments 3 and 4

			Unequal early				Unequal late			
	Experiment	Block	C	R	CO	RO	C	R	CO	RO
Overall	3	1	.64	.26	.06	.04	.08	.10	.42	.40
		2	.74	.20	.03	.03	.04	.02	.52	.42
		3	.74	.24	.02	0	0	.02	.56	.42
		4	.72	.26	.02	0	.01	.02	.58	.39
	4	4	.72	.26	.01	0	.02	.01	.44	.52
Early	3	1	.79	.20	0	.01	.04	.10	.49	.38
		2	.81	.18	.01	0	.01	0	.58	.40
		3	.79	.20	.01	0	.01	.01	.60	.38
		4	.80	.20	0	0	.01	.01	.57	.40
	4	4	.76	.23	0	.01	.01	.01	.40	.58
Late	3	1	.54	.30	.11	.05	.12	.11	.37	.44
		2	.69	.22	.04	.04	.05	.04	.47	.44
		3	.71	.26	.02	.01	0	.03	.53	.45
		4	.66	.30	.04	0	.01	.03	.59	.38
	4	4	.67	.31	.02	0	.05	.01	.50	.44

Table 55. Complete data for test item PC-Experiments 3 and 4

			Unequal early				Unequal late			
	Experiment	Block	C	R	CO	RO	C	R	CO	RO
Overall	3	1	.66	.17	.11	.05	.08	.14	.58	.20
		2	.91	.05	0	.04	0	.05	.84	.10
		3	.94	.02	0	.03	.02	.01	.94	.03
		4	.93	.05	0	.02	0	0	.96	.03
	4	4	.96	.04	0	0	0	0	.95	.05
Early	3	1	.81	.13	.05	.01	.01	.07	.68	.24
		2	.98	.02	0	0	0	.01	.90	.08
		3	.96	.02	0	.01	.04	0	.94	.02
		4	.96	.04	0	0	0	0	.99	.01
	4	4	.95	.05	0	0	0	0	.99	.01
Late	3	1	.55	.20	.16	.08	.12	.19	.51	.18
		2	.86	.06	.01	.07	.01	.08	.80	.12
		3	.93	.02	.01	.04	0	.02	.95	.04
		4	.91	.05	0	.04	0	.01	.95	.04
	4	4	.96	.04	0	0	0	0	.89	.11

Table 56. Complete data for test item PR-Experiments 3 and 4

			Unequal early				Unequal late			
	Experiment	Block	C	R	CO	RO	C	R	CO	RO
Overall	3	1	.13	.65	.16	.05	.03	.10	.24	.63
		2	.05	.87	.07	0	.02	.03	.10	.86
		3	.02	.95	.03	0	.03	.02	.03	.93
		4	0	.97	.02	0	.03	0	.04	.93
	4	4	.03	.96	.01	0	.01	0	.02	.97
Early	3	1	.07	.83	.08	.01	.01	.07	.14	.77
		2	0	1.00	0	0	.01	0	.01	.98
		3	.01	.98	.01	0	0	.01	.02	.96
		4	0	.99	.01	0	0	.01	.01	.98
	4	4	.03	.97	0	0	.01	0	.01	.98
Late	3	1	.18	.52	.22	.08	.04	.12	.32	.53
		2	.09	.78	.12	.01	.03	.04	.16	.77
		3	.03	.94	.04	0	.04	.02	.04	.90
		4	.01	.96	.03	0	.04	0	.05	.90
	4	4	.04	.94	.02	0	.01	0	.04	.95

Table 57. Complete data for test item PC+PR-Experiments 3 and 4

			Unequal early				Unequal late			
	Experiment	Block	C	R	CO	RO	C	R	CO	RO
Overall	3	1	.44	.43	.10	.03	.02	.09	.45	.44
		2	.35	.64	.01	0	.02	.03	.50	.46
		3	.35	.65	0	0	0	0	.48	.52
		4	.35	.64	.01	0	.01	0	.48	.51
	4	4	.21	.78	.01	0	0	0	.45	.54
Early	3	1	.42	.51	.06	.01	0	.05	.38	.57
		2	.25	.75	0	0	.01	.01	.45	.52
		3	.25	.75	0	0	0	0	.43	.57
		4	.20	.80	0	0	0	0	.40	.60
	4	4	.19	.81	0	0	0	0	.45	.55
Late	3	1	.46	.37	.12	.04	.04	.12	.50	.34
		2	.42	.56	.02	0	.02	.04	.53	.42
		3	.42	.57	0	.01	.01	0	.52	.47
		4	.46	.52	.02	0	.02	0	.54	.45
	4	4	.24	.70	.02	0	.01	.01	.45	.53

Table 58. Complete data for test item I+PC+PR-Experiments 3 and 4

			Unequal early				Unequal late			
	Experiment	Block	C	R	CO	RO	C	R	CO	RO
Overall	3	1	.65	.30	.04	.02	.06	.07	.34	.53
		2	.61	.38	.02	0	0	.02	.43	.56
		3	.49	.51	0	0	.01	0	.38	.61
		4	.48	.50	.01	0	0	0	.44	.56
	4	4	.44	.54	.01	0	0	.01	.42	.56
Early	3	1	.64	.32	.02	.01	.01	.05	.26	.68
		2	.49	.50	.01	0	0	0	.42	.58
		3	.34	.66	0	0	.01	0	.31	.68
		4	.29	.71	0	0	0	0	.39	.61
	4	4	.40	.58	.01	.01	0	.01	.42	.58
Late	3	1	.66	.28	.04	.02	.10	.09	.40	.41
		2	.70	.29	.02	0	0	.03	.44	.54
		3	.60	.40	0	0	.01	0	.44	.55
		4	.63	.35	.02	0	0	0	.48	.52
	4	4	.51	.48	.01	0	.01	.02	.43	.54

Table 59. Complete data for test item I+PC-Experiments 3 and 4

			Unequal early				Unequal late			
	Experiment	Block	C	R	CO	RO	C	R	CO	RO
Overall	3	1	.81	.15	.03	.02	.06	.15	.59	.20
		2	.94	.02	.01	.03	0	.05	.85	.10
		3	.93	.06	0	0	0	.02	.96	.03
		4	.95	.05	0	0	0	.03	.93	.04
	4	4	.96	.04	0	0	0	0	.98	.02
Early	3	1	.96	.02	.01	0	.01	.10	.71	.18
		2	1.00	0	0	0	0	0	.95	.05
		3	.94	.05	0	.01	0	0	.99	.01
		4	.92	.08	0	0	0	.01	.98	.01
	4	4	.97	.03	0	0	0	0	.99	.01
Late	3	1	.69	.25	.04	.03	.10	.19	.49	.22
		2	.89	.04	.02	.05	0	.09	.78	.13
		3	.93	.07	0	0	0	.03	.94	.04
		4	.97	.02	.01	0	0	.04	.90	.06
	4	4	.93	.07	0	0	0	.01	.97	.02

Table 60. Complete data for test item I+PR-Experiments 3 and 4

			Unequal early				Unequal late			
	Experiment	Block	C	R	CO	RO	C	R	CO	RO
Overall	3	1	.37	.52	.06	.06	.03	.03	.26	.68
		2	.16	.79	.05	0	.02	0	.12	.86
		3	.09	.91	0	0	.01	0	.08	.91
		4	.08	.92	0	0	.01	0	.08	.91
	4	4	.07	.92	.01	0	0	0	.07	.93
Early	3	1	.38	.60	.01	.01	0	.02	.24	.74
		2	.10	.89	.01	0	.02	0	.05	.93
		3	.05	.95	0	0	0	0	.01	.99
		4	.05	.95	0	0	0	0	.02	.98
	4	4	.06	.94	0	0	0	0	.06	.94
Late	3	1	.36	.46	.09	.10	.05	.04	.28	.63
		2	.21	.70	.07	.01	.02	.01	.17	.80
		3	.12	.88	.01	0	.02	.01	.12	.85
	4	4	.10	.90	0	0	.02	0	.12	.87
		4	.09	.09	.02	0	0	.01	.09	.90

APPENDIX C

TABLES FOR EXPERIMENTS 5 AND 6

Table 61. Complete probability estimates for P(S| D)-Experiment 5

		Unequal early		Unequal late	
Block		C	R	C	R
1	Imperfect	83.43	52.70	67.86	78.44
	Perfect	83.56	80.04	76.95	80.52
	Imperfect	34.15	33.41	41.08	36.25
	Perfect	28.04	33.20	34.41	28.41
2	Imperfect	81.08	65.66	72.92	74.44
	Perfect	93.86	82.38	86.22	85.71
	Imperfect	16.82	27.51	20.76	23.51
	Perfect	14.38	22.39	22.72	19.84
3	Imperfect	83.18	77.30	76.03	73.23
	Perfect	93.57	90.90	89.77	88.25
	Imperfect	13.32	24.41	18.24	18.87
	Perfect	11.96	17.75	15.15	13.90
4	Imperfect	84.11	76.30	80.73	72.42
	Perfect	90.81	90.34	91.08	92.86
	Imperfect	15.82	22.58	18.75	11.76
	Perfect	13.90	13.84	15.14	13.13

Table 62. Complete probability estimates for P(S| D)-Experiment 6

		Unequal early		Unequal late	
Block		C	R	C	R
1	Imperfect	72.59	37.72	62.76	65.38
	Perfect	66.48	57.90	72.59	64.31
	Imperfect	28.28	41.17	37.97	34.28
	Perfect	22.93	5052	26.83	27.69
2	Imperfect	82.21	63.66	70.17	74.14
	Perfect	91.38	89.21	89.55	89.21
	Imperfect	15.66	27.97	28.55	22.03
	Perfect	10.55	21.93	9.97	17.62
3	Imperfect	82.31	76.10	72.35	70.72
	Perfect	87.14	93.14	93.00	92.62
	Imperfect	5.72	23.24	17.17	4.21
	Perfect	11.79	23.69	5.62	2.76
4	Imperfect	76.90	74.59	89.66	76.00
	Perfect	87.93	96.69	88.45	87.45
	Imperfect	12.48	15.28	12.62	9.24
	Perfect	15.55	14.17	5.90	0.38

Table 63. Complete probability estimates for P(D|S)-Experiment 5

Block		Unequal early		Unequal late	
		C	R	C	R
1	Imperfect	69.28	46.25	57.39	61.29
	Perfect	81.04	71.85	61.98	71.06
	Imperfect	25.72	35.19	35.58	36.53
	Perfect	20.22	29.61	27.81	22.15
2	Imperfect	68.66	55.48	61.33	65.43
	Perfect	90.48	79.34	80.27	83.47
	Imperfect	19.77	22.53	26.49	22.14
	Perfect	13.66	19.27	18.82	12.90
3	Imperfect	74.10	61.30	63.99	62.81
	Perfect	83.15	85.52	83.63	84.51
	Imperfect	13.35	25.57	18.53	17.87
	Perfect	8.37	11.95	17.44	12.43
4	Imperfect	69.75	57.73	63.97	61.77
	Perfect	90.41	85.41	84.79	85.60
	Imperfect	15.44	20.13	18.77	16.66
	Perfect	12.03	13.09	13.79	11.61

Table 64. Complete probability estimates for P(D|S)-Experiment 6

Block		Unequal early		Unequal late	
		C	R	C	R
1	Imperfect	61.69	32.31	58.83	53.38
	Perfect	68.17	63.86	72.83	66.24
	Imperfect	23.86	26.48	45.93	44.69
	Perfect	27.79	41.76	20.41	22.69
2	Imperfect	71.17	55.52	54.35	61.86
	Perfect	87.79	89.55	82.28	77.21
	Imperfect	16.31	26.17	17.55	13.66
	Perfect	22.35	14.83	9.48	11.24
3	Imperfect	66.00	48.86	55.41	61.62
	Perfect	82.28	86.79	87.76	93.83
	Imperfect	9.14	32.76	19.86	12.90
	Perfect	9.24	11.28	2.48	12.52
4	Imperfect	55.24	52.90	64.17	66.79
	Perfect	81.90	86.69	83.45	91.00
	Imperfect	4.66	16.93	6.90	3.35
	Perfect	12.14	11.10	6.90	3.66

Table 65. Complete probability estimates for P(S|D)- early versus late-Experiment 5

		Early learners				Late learners			
		Unequal early		Unequal late		Unequal early		Unequal late	
Block		C	R	C	R	C	R	C	R
1	Imperfect	89.96	54.00	71.54	81.61	77.84	52.47	63.12	73.53
	Perfect	86.70	88.91	79.28	90.30	78.53	66.97	72.88	67.41
	Imperfect	23.59	31.28	33.04	21.83	50.41	37.50	53.91	58.12
	Perfect	17.96	22.80	26.83	20.24	43.41	46.38	43.25	41.03
2	Imperfect	83.52	68.65	72.09	70.67	76.97	63.41	76.41	79.06
	Perfect	91.89	87.24	91.02	90.83	96.50	74.84	78.88	77.91
	Imperfect	13.59	19.35	14.91	16.52	22.00	40.10	28.25	34.28
	Perfect	12.50	16.30	15.35	15.59	17.53	28.72	30.91	26.56
3	Imperfect	80.46	79.30	78.06	73.50	86.56	76.84	73.91	72.00
	Perfect	89.41	90.63	89.28	83.87	99.34	91.00	90.16	94.19
	Imperfect	12.30	22.67	15.33	16.13	15.19	27.66	23.00	23.41
	Perfect	13.38	14.04	17.52	17.11	10.34	22.06	12.22	9.72
4	Imperfect	85.13	76.06	81.41	70.28	82.16	79.03	80.72	74.62
	Perfect	89.24	92.04	92.33	90.00	92.78	87.59	89.00	96.75
	Imperfect	15.50	23.11	15.89	11.48	16.78	22.53	21.88	12.53
	Perfect	13.85	12.98	12.15	14.59	14.41	13.94	19.91	11.44

Table 66. Complete probability estimates for P(S|D)-early versus late-Experiment 6

		Early learners				Late learners			
		Unequal early		Unequal late		Unequal early		Unequal late	
Block		C	R	C	R	C	R	C	R
1	Imperfect	69.21	34.58	62.73	68.42	79.00	43.70	62.80	59.60
	Perfect	64.26	64.21	80.52	68.95	70.70	45.90	57.50	55.50
	Imperfect	19.74	35.63	33.16	28.63	44.50	51.70	47.10	45.00
	Perfect	13.16	58.00	18.95	17.74	41.50	36.30	41.80	46.60
2	Imperfect	86.68	58.37	70.42	74.74	73.70	73.70	69.70	73.00
	Perfect	92.53	85.58	94.58	90.00	89.20	96.10	80.00	87.70
	Imperfect	13.58	22.53	21.00	10.68	19.60	38.30	42.90	43.60
	Perfect	6.05	22.84	.11	16.21	19.10	20.20	28.70	20.30
3	Imperfect	86.84	79.47	74.74	81.53	73.70	69.70	67.80	50.20
	Perfect	81.84	98.90	95.79	96.26	97.20	82.20	87.70	85.70
	Imperfect	4.42	16.37	13.21	0	8.20	36.30	24.70	12.20
	Perfect	13.16	23.95	7.42	3.32	9.20	23.20	2.20	1.70
4	Imperfect	76.74	72.95	88.84	72.47	77.20	77.70	91.20	82.70
	Perfect	90.95	99.84	94.11	92.05	82.20	90.70	77.70	78.70
	Imperfect	10.00	5.26	12.11	0	17.20	34.30	13.60	26.80
	Perfect	15.79	13.16	5.26	0	15.10	16.10	7.10	1.10

Table 67. Complete probability estimates for P(D|S)-early versus late-Experiment 5

		Early learners				Late learners			
		Unequal early		Unequal late		Unequal early		Unequal late	
Block		C	R	C	R	C	R	C	R
1	Imperfect	70.46	45.09	57.85	62.96	66.62	49.38	56.97	59.25
	Perfect	82.37	80.85	71.17	76.98	78.53	58.03	47.56	63.22
	Imperfect	19.02	36.41	31.61	24.13	36.16	34.53	40.84	55.50
	Perfect	17.28	18.56	16.28	19.56	25.06	44.84	45.25	26.56
2	Imperfect	69.89	56.59	58.24	64.41	65.91	55.62	66.12	65.81
	Perfect	86.63	86.48	87.00	87.72	95.72	68.44	69.97	76.84
	Imperfect	10.65	20.63	22.65	8.72	33.50	25.97	32.84	42.12
	Perfect	11.17	12.11	12.06	8.96	17.66	30.16	29.12	18.97
3	Imperfect	70.74	59.36	64.61	61.91	78.12	66.06	65.09	62.94
	Perfect	81.09	87.38	82.02	82.02	85.59	82.44	85.44	87.59
	Imperfect	13.02	22.22	13.17	10.59	14.25	31.19	26.81	28.91
	Perfect	7.39	7.89	12.65	13.28	10.03	16.59	24.88	11.59
4	Imperfect	67.39	56.52	63.91	58.61	72.19	61.28	66.06	65.12
	Perfect	86.39	83.13	86.61	83.15	95.88	88.22	84.50	88.66
	Imperfect	9.89	16.83	16.72	11.41	23.91	25.20	22.31	24.72
	Perfect	8.35	11.30	10.06	11.00	17.69	16.06	19.56	12.84

Table 68. Complete probability estimates for P(D|S)- early versus late-Experiment 6

		Early learners				Late learners			
		Unequal early		Unequal late		Unequal early		Unequal late	
Block		C	R	C	R	C	R	C	R
1	Imperfect	66.58	34.68	54.32	53.79	52.40	27.80	67.40	52.60
	Perfect	72.68	56.58	73.73	62.16	59.60	77.70	71.10	74.00
	Imperfect	16.05	17.63	40.58	33.16	38.70	43.30	56.10	66.60
	Perfect	13.42	27.84	13.68	18.42	55.10	68.20	33.20	38.80
2	Imperfect	70.63	50.42	55.11	63.42	72.20	65.20	52.90	58.90
	Perfect	84.16	91.05	83.90	87.00	94.70	86.70	79.20	58.60
	Imperfect	12.16	20.37	2.90	6.32	24.20	37.20	45.40	27.60
	Perfect	6.11	9.21	5.26	3.32	24.20	25.50	17.50	26.30
3	Imperfect	70.95	37.21	47.90	61.05	56.60	56.50	69.70	62.70
	Perfect	82.00	98.95	92.63	93.11	82.80	63.70	78.50	95.20
	Imperfect	5.21	31.00	18.95	5.68	16.60	36.10	21.60	26.60
	Perfect	5.84	16.37	3.21	5.90	15.70	1.60	1.10	25.10
4	Imperfect	54.84	49.90	59.21	61.37	56.00	58.60	73.60	77.10
	Perfect	83.16	92.58	84.74	92.47	79.50	75.50	81.00	88.20
	Imperfect	5.47	11.84	8.90	2.95	3.10	26.60	3.10	4.10
	Perfect	14.26	6.37	9.95	2.63	8.10	20.10	1.10	5.60

Table 69. Complete data for test item I-Experiment 5

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.61	.21	.13	.05	.29	.54	.10	.08
	2	.66	.21	.09	.04	.35	.56	.03	.05
	3	.70	.25	.03	.01	.41	.54	.01	.03
	4	.67	.26	.04	.03	.40	.53	.03	.05
Early	1	.73	.14	.09	.04	.32	.62	.02	.04
	2	.76	.16	.05	.02	.36	.61	.01	.02
	3	.74	.23	.01	.02	.42	.55	0	.02
	4	.70	.24	.02	.05	.42	.55	0	.02
Late	1	.44	.31	.19	.06	.25	.41	.20	.14
	2	.52	.28	.14	.06	.34	.50	.06	.09
	3	.66	.28	.06	0	.39	.53	.03	.05
	4	.64	.30	.06	0	.36	.48	.06	.09

Table 70. Complete data for test item I-Experiment 6

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.62	.28	.03	.07	.41	.33	.10	.16
	2	.74	.26	0	0	.33	.53	.03	.10
	3	.71	.28	.02	0	.47	.52	.02	0
	4	.67	.33	0	0	.40	.60	0	0
Early	1	.68	.26	0	.05	.50	.34	.03	.13
	2	.72	.29	0	0	.37	.58	0	.05
	3	.66	.32	.03	0	.37	.61	.03	0
	4	.74	.26	0	0	.37	.63	0	0
Late	1	.50	.30	.10	.10	.25	.30	.25	.20
	2	.80	.20	0	0	.25	.45	.10	.20
	3	.80	.20	0	0	.65	.35	0	0
	4	.55	.45	0	0	.45	.55	0	0

Table 71. Complete data for test item PC-Experiment 5

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.78	.13	.06	.03	.58	.16	.11	.15
	2	.90	.05	.02	.03	.78	.08	.05	.09
	3	.91	.06	0	.03	.87	.03	.03	.07
	4	.95	.03	.01	.01	.94	.01	0	.05
Early	1	.86	.08	.03	.03	.72	.06	.08	.14
	2	.94	.03	.02	.01	.90	.03	.03	.03
	3	.92	.05	0	.02	.95	0	.02	.03
	4	.97	.03	0	0	.99	0	0	.01
Late	1	.67	.20	.11	.02	.39	.30	.16	.16
	2	.86	.08	.02	.05	.59	.16	.08	.17
	3	.89	.06	0	.05	.77	.18	.03	.12
	4	.92	.03	.03	.01	.88	.02	0	.11

Table 72. Complete data for test item PC-Experiment 6

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.66	.19	.05	.10	.53	.28	.05	.14
	2	.81	.17	0	.02	.91	.03	.02	.03
	3	.83	.12	.05	0	.98	0	0	.02
	4	.90	.03	.05	.02	.97	.03	0	0
Early	1	.71	.13	.08	.08	.61	.26	.03	.11
	2	.79	.18	0	.03	.92	.03	0	.05
	3	.87	.08	.05	0	1.00	0	0	0
	4	.87	.05	.05	.03	.97	.03	0	0
Late	1	.55	.30	0	.15	.40	.30	.10	.20
	2	.85	.15	0	0	.90	.05	.05	0
	3	.75	.20	.05	0	.95	0	0	.05
	4	.95	0	.05	0	.95	.05	0	0

Table 73. Complete data for test item PR-Experiment 5

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.15	.64	.17	.05	.21	.63	.05	.11
	2	.08	.83	.06	.03	.08	.84	.05	.03
	3	.03	.89	.07	.01	.04	.92	.03	.03
	4	.01	.92	.05	.02	.03	.91	.03	.03
Early	1	.04	.84	.09	.03	.14	.79	.03	.03
	2	.01	.96	.03	0	.04	.96	0	0
	3	0	.94	.06	0	0	.99	0	.01
	4	0	.97	.02	.01	.01	.97	.02	0
Late	1	.30	.36	.28	.06	.31	.39	.08	.22
	2	.17	.66	.11	.06	.12	.67	.12	.08
	3	.06	.83	.08	.03	.09	.83	.03	.05
	4	.03	.84	.09	.03	.06	.83	.05	.06

Table 74. Complete data for test item PR-Experiment 6

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.17	.43	.29	.10	.29	.5	.02	.14
	2	.09	.84	.03	.03	.14	.86	0	0
	3	0	.97	0	.03	.05	.95	0	0
	4	.02	.95	.03	0	.03	.97	0	0
Early	1	.13	.63	.16	.08	.26	.63	.03	.08
	2	.13	.87	0	0	.13	.87	0	0
	3	0	1.00	0	0	.05	.95	0	0
	4	.03	.95	.03	0	0	1.00	0	0
Late	1	.25	.05	.55	.15	.35	.40	0	.25
	2	0	.80	.10	.10	.15	.85	0	0
	3	0	.90	0	.10	.05	.95	0	0
	4	0	.95	.05	0	.10	.90	0	0

Table 75. Complete data for test item PC+PR-Experiment 5

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.38	.51	.08	.03	.46	.39	.03	.12
	2	.37	.58	.02	.03	.46	.47	.02	.05
	3	.36	.61	.03	.01	.43	.50	.03	.03
	4	.35	.60	.04	.01	.46	.51	0	.03
Early	1	.30	.62	.05	.02	.48	.40	.03	.09
	2	.32	.67	0	.01	.49	.47	.02	.02
	3	.29	.68	.01	.01	.49	.46	.03	.02
	4	.30	.68	.01	0	.53	.46	0	.01
Late	1	.48	.36	.11	.05	.44	.38	.03	.16
	2	.45	.45	.05	.05	.41	.48	.02	.09
	3	.45	.50	.05	0	.34	.58	.03	.05
	4	.42	.48	.08	.02	.34	.59	0	.06

Table 76. Complete data for test item PC+PR-Experiment 6

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.33	.45	.17	.05	.43	.48	0	.09
	2	.38	.59	.03	0	.48	.50	0	.02
	3	.31	.69	0	0	.50	.50	0	0
	4	.36	.64	0	0	.50	.48	.02	0
Early	1	.37	.47	.13	.03	.45	.50	0	.05
	2	.37	.63	0	0	.50	.50	0	0
	3	.31	.68	0	0	.50	.50	0	0
	4	.34	.66	0	0	.50	.50	0	0
Late	1	.25	.40	.25	.10	.40	.45	0	.15
	2	.40	.50	.10	0	.45	.50	0	.05
	3	.30	.70	0	0	.50	.50	0	0
	4	.40	.60	0	0	.50	.50	.05	0

Table 77. Complete data for test item I+PC+PR-Experiment 5

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.50	.38	.07	.04	.38	.50	.04	.08
	2	.57	.40	.01	.01	.39	.54	.03	.04
	3	.49	.49	.01	0	.47	.48	.02	.03
	4	.44	.52	.03	.01	.49	.47	.01	.03
Early	1	.46	.46	.03	.06	.43	.50	.02	.05
	2	.52	.46	.01	0	.45	.52	.01	.03
	3	.43	.55	.02	0	.55	.42	.01	.02
	4	.39	.59	.01	.01	.53	.44	.01	.02
Late	1	.59	.26	.14	.02	.29	.5	.09	.13
	2	.64	.31	.02	.03	.30	.59	.05	.05
	3	.59	.41	.01	0	.36	.56	.05	.03
	4	.51	.41	.05	.02	.43	.52	.02	.04

Table 78. Complete data for test item I+PC+PR-Experiment 6

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.66	.24	.08	.03	.52	.38	.05	.05
	2	.63	.37	0	0	.47	.49	.01	.03
	3	.49	.51	0	0	.51	.47	0	.02
	4	.42	.58	0	0	.47	.52	0	.02
Early	1	.70	.26	.04	0	.59	.37	0	.04
	2	.68	.32	0	0	.47	.53	0	0
	3	.53	.47	0	0	.61	.39	0	0
	4	.42	.58	0	0	.54	.45	0	.01
Late	1	.58	.20	.15	.08	.38	.40	.15	.08
	2	.52	.48	0	0	.48	.42	.02	.08
	3	.42	.58	0	0	.32	.62	0	.05
	4	.42	.58	0	0	.32	.65	0	.02

Table 79. Complete data for test item I+PC*-Experiment 5

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.27	.26	.39	.08	.39	.27	.17	.16
	2	.25	.19	.50	.06	.40	.18	.18	.23
	3	.27	.16	.54	.03	.48	.15	.22	.16
	4	.24	.12	.61	.02	.50	.16	.13	.21
Early	1	.29	.20	.44	.08	.37	.33	.13	.17
	2	.23	.14	.60	.03	.38	.22	.21	.20
	3	.23	.14	.62	.01	.47	.14	.24	.15
	4	.25	.08	.66	.01	.53	.17	.14	.15
Late	1	.20	.36	.34	.09	.42	.20	.23	.14
	2	.28	.27	.36	.09	.45	.12	.16	.27
	3	.30	.19	.45	.06	.50	.14	.19	.17
	4	.22	.19	.55	.05	.47	.12	.12	.28

Table 80. Complete data for test item I+PR*-Experiment 5

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.30	.16	.13	.41	.16	.40	.25	.19
	2	.24	.11	.10	.55	.13	.47	.21	.19
	3	.25	.08	.06	.61	.12	.55	.22	.11
	4	.24	.11	.02	.63	.12	.52	.21	.15
Early	1	.30	.11	.11	.49	.09	.52	.21	.18
	2	.26	.06	.02	.65	.11	.56	.16	.16
	3	.24	.09	.03	.64	.14	.59	.21	.06
	4	.25	.08	0	.67	.12	.61	.15	.12
Late	1	.30	.23	.16	.31	.23	.23	.33	.20
	2	.19	.17	.22	.42	.14	.34	.28	.23
	3	.23	.08	.11	.58	.09	.52	.25	.14
	4	.20	.16	.05	.59	.12	.42	.30	.16

Table 81. Complete data for test item PC+PR*-Experiment 5

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.41	.15	.10	.34	.16	.34	.35	.15
	2	.48	.07	.06	.40	.11	.39	.46	.04
	3	.46	.04	.01	.48	.11	.50	.39	.01
	4	.47	.06	.02	.44	.11	.50	.38	.01
Early	1	.41	.11	.08	.40	.14	.41	.34	.11
	2	.47	.04	.02	.47	.14	.36	.49	.01
	3	.46	.03	0	.51	.10	.50	.40	0
	4	.46	.03	.01	.50	.12	.46	.41	.01
Late	1	.42	.19	.14	.25	.17	.23	.38	.22
	2	.47	.11	.11	.31	.06	.45	.41	.08
	3	.47	.06	.03	.44	.12	.52	.34	.02
	4	.48	.11	.05	.36	.09	.58	.31	.02

Table 82. Complete data for test item I+PC+PR*-Experiment 5

		Unequal early				Unequal late			
	Block	C	R	CO	RO	C	R	CO	RO
Overall	1	.56	.13	.10	.21	.12	.22	.52	.14
	2	.71	.09	.03	.17	.05	.23	.64	.08
	3	.67	.06	.02	.25	.04	.28	.65	.03
	4	.72	.04	.02	.22	.04	.20	.71	.04
Early	1	.66	.08	.03	.23	.08	.23	.63	.06
	2	.75	.05	.01	.18	.02	.25	.70	.03
	3	.73	.02	.02	.23	.01	.23	.75	.01
	4	.74	.02	.01	.23	.02	.18	.77	.02
Late	1	.41	.22	.19	.19	.17	.20	.38	.25
	2	.64	.14	.06	.16	.09	.22	.55	.14
	3	.58	.12	.02	.28	.08	.36	.50	.06
	4	.69	.08	.03	.20	.08	.23	.61	.08

BIBLIOGRAPHY

Binder, A., & Estes, W.K. (1966). Transfer of response in visual recognition situations as a function of frequency variables. Psychological Monographs, 80,(23,Whole No.631).

Broadbent, D., A question of levels: Comment on McClelland and Rumelhart. Journal of Experimental Psychology: General, 114, 189-192.

Carroll, J. S., & Siegler, R. S. (1977). Strategies for the use of base-rate information. Organizational and Human Performance, 19, 392-402.

Christensen-Szalanski, J. J. J. & Beach, L. R. (1982). Experience and the base-rate fallacy. Organizational Behavior and Human Performance, 29, 270-278.

Christensen-Szalanski, J. J. J. & Bushyhead, J. B. (1981). Physicians' use of probabilistic information in a real clinical setting. Journal of Experimental Psychology: Human Perception and Performance, 7, 928-935.

Chumbley, J.I. (1996). CLEM: Concept learning by exemplar memorization. Unpublished manuscript.

Erikson, M. A., & Kruschke, J.K. (1998). Rules and exemplars in category learning. Journal of Experimental Psychology: General, 127, 107-140.

Estes, W. K. (1993). Models of Categorization and Category Learning. The Psychology of Learning and Motivation, 29, 15-56.

Estes, W. K. (1994). Classification and Cognition: Oxford Psychological Series #22. New York, Clarendon Press.

Estes, W. K., Campbell, J. A., Hatsopoulos, N., & Hurwitz, J. B. (1989). Base-rate effects in category learning: A comparison of parallel network and memory storage retrieval models. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 556-571.

Gluck, M. A. (1992). Stimulus sampling and distributed representations in adaptive network theories of learning. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), Essays in honor of William K. Estes (Vol. 1, pp. 169-199). Hillsdale, NJ: Erlbaum.

Gluck, M. A., & Bower, G. H. (1988a). Evaluating an adaptive network model of human learning. Journal of Memory and Language, 27, 166-195.

Gluck, M. A., & Bower, G. H.(1988b). From conditioning to category learning: An adaptive network model. Journal of Experimental Psychology: General, 117, 227-247.

Gluck, M. A., & Bower, G. H. (1990). Component and pattern information in adaptive networks. Journal of Experimental Psychology: General, 119, 105-109.

Hansen, McCann, & Myers, J. L. (1985). Rote versus conceptual emphases in teaching elementary probability. Journal for Research in Mathematics Education, 16, 364-374.

Kahneman, D., & Tversky, A.(1973). On the psychology of prediction. Psychological Review, 80, 237-251.

Koehler, J. J. (1996). The base-rate fallacy reconsidered: Descriptive, normative, and methodological challenges. Behavioral and Brain Sciences, 19, 1-53.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. Psychological Review, 99, 22-44.

Kruschke, J. K. (1996). Base rates in category learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 3-26.

Kruschke, J. K., & Erikson, M. A. (1995). Five principles for models of category learning. [On-line] Unpublished Manuscript. Available: World Wide Web URL:http://www.indiana.edu/ ~kruschke/fiveprinc abstract.html.

Lewandowsky, S. (1995). Base-rate neglect in ALCOVE: A critical reevaluation. Psychological Review, 102, 185-191.

Luce, R.D. (1963). Detection and recognition. In R.D. Luce, R. R. Bush, and E. Galanter (Eds.), Handbook of Mathematical Psychology (pp. 103-189). New York: Wiley.

Markman, A. B. (1989) LMS rules and the inverse base-rate effect: Comment on Gluck and Bower(1988). Journal of Experimental Psychology: General, 118, 417-421.

Medin, D. L. & Bettger J. G. (1991). Sensitivity to Changes in Base-rate Information. American Journal of Psychology, 104, 311-332.

Medin, D. L., & Edelson, S. M. (1988). Problem structure and the use of base-rate information from experience. Journal of Experimental Psychology: General, 117, 68-85.

Medin, D. L., & Florian, J. E.(1992). Abstraction and selective coding in exemplar-based models of categorization. In A. Healy, S. Kosslyn, and R. Shiffrin (Eds.), Essays in Honor of William K. Estes. (pp. 207-234). Hillsdale, New Jersey: Erlbaum.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. Psychological Review, 85, 207-238.

Myers, J.L. (1982). The role of explanation in teaching elementary probability. Proceedings of the First International Conference on Teaching Statistics, 11, 802-812.

Myers, J. L., Lohmeier, J. H., & Well, A. D. (1994). Modeling probabilistic categorization data: Exemplar memory and connectionist nets. Psychological Science, 5, 83-89.

Nelson, M. W. (1996). Context and the inverse base-rate effect, Journal of Behavioral Decision Making, 9, 23-40.

Nosofsky, R. M., Kruschke, J. K., & Mc Kinley, S. C. (1992). Combining exemplar-based category representations and connectionist learning rules. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 211-233.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W.F. Prokasy (Eds.), Classical conditioning: II. Current research and theory. (pp. 64-99). New York: Appelton-Century-Crofts.

Shanks, D. R. (1995). Problems in the Behavioral Sciences 13. The Psychology of Associative Learning. Cambridge: Cambridge University Press.

Shanks, D. R. (1992). Connectionists accounts of the inverse base-rate effect in categorization. Connection Science, 4, 3-18.

Wickens, T.D. (1989). Multiway contingency tables analysis for the social sciences. Hillsdale, NJ: Erlbaum.

Widrow, B., & Hoff, M. E. (1960). Adaptive switching circuits. Institute of Radio Engineers, Western Electronics Show and Convention, Convention Record, 4, 96-194.